

Bibliotecas y publicaciones digitales

JUAN VOUTSSÁS MÁRQUEZ



La presente obra está bajo una licencia de:

http://creativecommons.org/licenses/by-nc-sa/3.0/deed.es_MX



Atribución-No Comercial-Licenciamiento Recíproco 3.0 Unported

Eres libre de:



copiar, distribuir y comunicar públicamente la obra



hacer obras derivadas

Bajo las condiciones siguientes:



Atribución — Debes reconocer la autoría de la obra en los términos especificados por el propio autor o licenciante.



No comercial — No puedes utilizar esta obra para fines comerciales.



Licenciamiento Recíproco — Si alteras, transformas o creas una obra a partir de esta obra, solo podrás distribuir la obra resultante bajo una licencia igual a ésta.

Esto es un resumen fácilmente legible del:

[texto legal \(de la licencia completa\)](#)

En los casos que sea usada la presente obra, deben respetarse los términos especificados en esta licencia.



Bibliotecas y publicaciones digitales

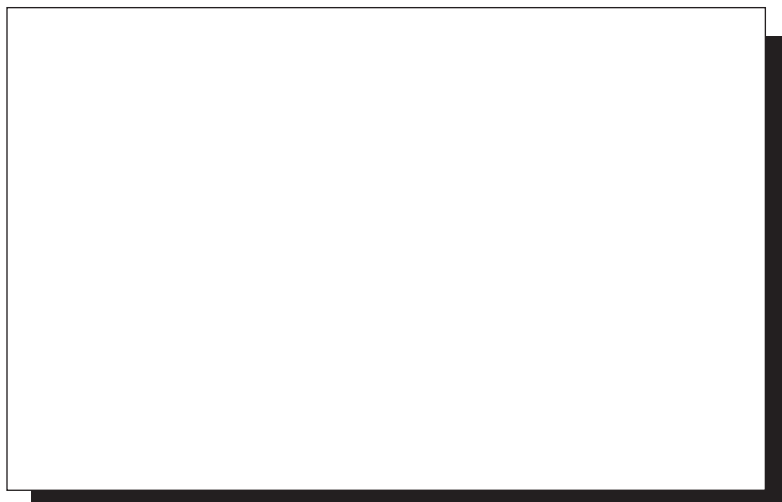
COLECCIÓN
TECNOLOGÍAS DE LA INFORMACIÓN
Centro Universitario de Investigaciones Bibliotecológicas

Bibliotecas y publicaciones digitales

Juan Voutssás Márquez



Universidad Nacional Autónoma de México
2006



Diseño de portada: Mario Ocampo Chávez

Primera Edición 2006

DR © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad Universitaria, 04510, México, D. F.

Impreso y hecho en México

ISBN:



*"Si no sabes
hacia dónde
vas,
cualquier
camino
sirve".*

**Ilustración de Sir John Tenniel,
para la primera edición de Lewis Carroll, de 1865.*

Tabla de contenido

INTRODUCCIÓN

xiii

Capítulo 1 LA BIBLIOTECA DIGITAL

1.1 Concepto	
1.2 Definición	
1.3 Problemática	
1.4 Enfoques	

Capítulo 2 COLECCIONES DIGITALES CONVERTIDAS

2.1 Desarrollo de colecciones: concepto y antecedentes	
2.2 Conversión de documentos originalmente en soportes tradicionales	
2.2.1 Conservación y preservación	
2.3 Criterios para digitalizar	
2.4 Formatos digitales	
2.4.1 Texto	
2.4.2 Imagen fija	
2.4.3 Audio	
2.4.4 Imagen dinámica	
2.5 Compresión y almacenamiento	

Capítulo 3

COLECCIONES DE DOCUMENTOS ORIGINALMENTE DIGITALES

3.1 Publicación electrónica o digital	
3.2 Revistas científicas	
3.3 Libro electrónico	
3.3.1 Dispositivos	
3.3.2 Seguridad.....	
3.3.3 Formatos	
3.4 Otras publicaciones electrónicas	
3.5 Hipervínculos a otras colecciones	

Capítulo 4

HERRAMIENTAS PARA EL REGISTRO DE DOCUMENTOS DIGITALES

4.1 Del registro documental manual al electrónico	
4.2 Metadatos	
4.3 Formatos de máquina y lenguajes de marcado	
4.3.1 MARC.....	
4.3.2 Metalenguajes	
4.3.3 HTML.....	
4.3.4 SGML	
4.3.5 XML.....	

Capítulo 5

PUBLICACIONES ELECTRÓNICAS Y DERECHOS DE AUTOR Y DE PROPIEDAD

5.1 Derecho de autor y derecho de copia	
5.1.1 La Ley Federal del Derecho de Autor en México	
5.1.2 La relación hombre-libro y el equilibrio entre editores y usuarios	

5.2 Las iniciativas internacionales sobre los derechos	
5.2.1 La iniciativa europea	
5.2.2 El “ <i>fair use</i> ” o “ <i>uso ético</i> ” de la información	
5.2.3 El “ <i>copyleft</i> ” y “ <i>Creative Commons</i> ”	

Capítulo 6

EL BIBLIOTECARIO PROFESIONAL DIGITAL

6.1 Perfil del bibliotecario para la biblioteca digital	
6.2 El bibliotecario digital en Norteamérica	
6.3 El bibliotecario digital en Europa	
6.4 El bibliotecario digital en América Latina	

RESUMEN.....	
CONCLUSIONES.....	253

GLOSARIO.....	275
REFERENCIASBIBLIOGRÁFICAS.....	283
ANEXOS	307
ANEXO 1.- Producción mundial estimada de contenidos originales digitales	308
ANEXO 2.- Proceso evolutivo de Internet en la PyME.....	309
ANEXO 3.- México - Numeralia	310
ANEXO 4.- Tablas de datos y formatos utilizados en proyectos de digitalización	311

*“El bien de un libro consiste en ser leído.
Un libro está hecho de signos que hablan
de otros signos que, a su vez, hablan de las cosas.
Sin unos ojos que lo lean, un libro contiene signos
que no producen conceptos. Y por tanto, es mudo.”*

*El Nombre de la Rosa,
de Umberto Eco, p. 482.*

INTRODUCCIÓN

Casi desde el principio de mi proyecto de investigación doctoral, me percaté de que en el intento de desarrollar un modelo de bibliotecas para nuestro país, había dos aproximaciones básicas al mismo, las cuales eran sustancialmente distintas y mutuamente excluyentes una de la otra: por un lado, existía la posibilidad de estudiar y explicar a la biblioteca digital desde un enfoque reduccionista, como un ente aislado o cuando mucho, como un grupo de ellas dentro de una organización. Por el otro lado podía analizarse bajo un enfoque holístico y bajo éste debería verse a un modelo de bibliotecas digitales como el total de ellas, un inmenso conjunto o sistema, o una enorme organización a nivel de todo el país, y por tanto sus partes serían sus sectores, componentes y relaciones funcionales a nivel nacional.

El tratamiento y objeto de estudio eran por consecuencia distintos. Bajo el primer enfoque, el objeto de estudio es la biblioteca digital en sí misma, sus partes y sus interrelaciones: las preguntas pertinentes al respecto serían entonces ¿qué es exactamente una biblioteca digital y qué rasgos la definen? ¿de dónde viene? ¿existe un sólo tipo de ellas o cuáles son sus variantes? ¿qué partes la conforman? ¿cuáles son sus servicios y colecciones y cómo se desarrollan ambos? ¿cómo es el mundo de las publicaciones digitales que se integran a sus colecciones? ¿cuáles son las semejanzas y las diferencias entre ellas y una biblioteca tradicional? ¿cómo se diseña y construye una biblioteca digital? ¿qué esperan los usuarios de una biblioteca de este tipo?

Éstas, entre otras preguntas que me hacía personalmente, son las que se hacen actualmente los encargados de las bibliotecas, los profesionales de las mismas, los profesores y los estudiantes de bibliotecología.

Entonces, abordar la investigación bajo este enfoque permitiría, teniendo como centro la biblioteca digital, poder analizar minuciosamente todas y cada una de sus partes principales, explicar sus interrelaciones, descubrir sus características, separar sus mitos y realidades y, enfocándose a un lector que fuese un profesional de la biblioteca, ayudarme y ayudarle a comprender de la manera más precisa posible el origen, evolución, estado del arte y tendencias de la biblioteca digital; y también informarle cómo concebir, diseñar y construir una biblioteca digital, disectando cada una de sus partes y sus interrelaciones; cómo se va construyendo una colección digital, con material digital original o digitalizado; explicarle cuándo, cómo y por qué debe digitalizarse tal o cual material y las ventajas y desventajas de hacerlo; cuáles implicaciones legales tiene el uso o copia de los materiales digitales y cómo el bibliotecario actual puede desenvolverse adecuadamente en ese entorno.

Por supuesto, este enfoque era sin duda atractivo. Parecía muy interesante poder adentrarse en detalle en las entrañas de la biblioteca digital, disectarla, analizarla y descubrir sus secretos y paradigmas, y tratar de explicárselos a otros bibliotecólogos que tienen la necesidad, el encargo, o simplemente la inquietud de explorar el mundo de la biblioteca digital. Naturalmente, habría que decantar algunas características propias a nuestro medio mexicano, teniendo en cuenta sus necesidades, recursos y contexto.

Esta aproximación no cuestiona en sí la naturaleza social de la biblioteca, ni la necesidad o pertenencia de su existencia. Éstas en sí ya están dadas por cerca de tres milenios de historia. Nace del supuesto de que un profesional de la biblioteca desea diseñar y construir una biblioteca digital pública, o para una escuela, un centro de investigación, una entidad gubernamental o una empresa privada; y podría tratar de explicar cómo hacerlo partiendo del entendimiento cabal de la misma. No por ser reduccionista y puntual este enfoque dejaba de ser atractivo como objeto de estudio.

La segunda aproximación o enfoque, como ya se ha mencionado, era el holístico. En este caso el objeto de estudio son *todas* las posibles bibliotecas digitales existentes o potencialmente creables en nuestro país. Las preguntas cambian: en primer lugar, es necesario replantear el problema desde la función misma de las bibliotecas a nivel nacional.

Para ello era indispensable abrir mucho más el horizonte y comenzar con un objeto de estudio mucho más amplio. De ahí, y dado que la biblioteca es una institución de servicio para la sociedad, es necesario entonces analizar detalladamente esa sociedad en la que se inserta; cuáles son sus características, sus necesidades y su problemática; por supuesto, hablamos de la sociedad mexicana actual. Pero esta sociedad no es una isla en el planeta; es necesario entonces analizar antes a la sociedad mundial para contextualizar dentro de ella a la sociedad mexicana. Se dice que actualmente la sociedad en la que vivimos a nivel mundial es la denominada *sociedad de la información*. Era necesario entonces empezar a estudiar y definir a esta así llamada sociedad actual: ¿qué rasgos la definen? ¿en qué es nueva y en qué es igual a otras predecesoras? ¿desde cuándo existe? ¿cuál es su contexto? ¿cuáles son sus tendencias, fortalezas y problemática? ¿cuáles son sus riesgos, injusticias y omisiones?

A continuación surgió el hecho donde existen múltiples proyectos de muchas naciones o conglomerados de ellas tendientes a la construcción de esa *sociedad de la información*, todos ellos de creación reciente, inclusive hay un proyecto mexicano: *e-México*. ¿Por qué se generaron todos esos proyectos? ¿cuáles son sus propuestas, tendencias, características, semejanzas y diferencias? ¿por qué tanta preocupación a nivel mundial de tenerlos? ¿por qué uno en México? ¿cómo puede ayudar a nuestra sociedad mexicana? ¿qué riesgos corremos de no hacerlo? y lo más importante, ¿qué responsabilidad tienen las bibliotecas y los bibliotecarios mexicanos con respecto a ellos?

De este estudio, saltó a la vista el hecho de que, independientemente que la biblioteca tiene múltiples funciones sociales, el puente primigenio entre las bibliotecas y la sociedad ha sido y es la educación, y más recientemente, el trabajo. Surgieron entonces nuevas preguntas a reponder: ¿cómo es la educación dentro de esta *sociedad de la información*? ¿en qué se parece o diferencia de la anterior? ¿está cambiando esa educación y hacia dónde? ¿cuáles son las nuevas modalidades educativas? y más importante aún, ¿cuáles son los nuevos modelos educativos? ¿cómo son los nuevos educadores y educandos en este contexto? Si están cambiando modalidades y modelos educativos y las bibliotecas están intrínsecamente ligadas a la educación, ¿se seguirán requiriendo

bibliotecas en un mundo *internetizado*? Si ése es el caso ¿cómo deben evolucionar las bibliotecas, digitales por supuesto, frente a este cambio educativo? ¿cuál es el perfil de los profesionales de la bibliotecología que las conciben, construyen y operan?

Estos dos últimos bloques de preguntas, como puede observarse, plantean una nueva relación *sociedad-educación-bibliotecas-trabajo*, que debe ser analizada, discutida y explicada con mucho cuidado: el futuro está en juego; es particularmente relevante el estudio de esta relación en nuestro país y, por tanto, nuestro futuro.

Una vez definido este contexto previo, podía pasarse a la elaboración de un modelo funcional nacional de bibliotecas digitales: ¿cómo las organizamos a nivel nacional? ¿partimos de una gran super-biblioteca digital? ¿será un selecto conjunto de ellas? ¿será todo un conglomerado articulado a todo lo largo del país? ¿cómo se concibe y organiza la propuesta? ¿quiénes deben intervenir en ella? ¿cuáles son las responsabilidades de los sectores nacionales en este proyecto? Bajo este enfoque los lectores serían los reponsables e interesados para planear los diversos sectores del país, y quienes podrían obener material interesante de esta propuesta para una mejor toma de decisiones.

Este enfoque, como puede observarse, y las preguntas que se plantean y a las que se pretendería responder son también a todas luces interesantes. Durante un tiempo existió la duda acerca de seleccionar alguno de los dos enfoques y se encontraron válidas razones para desarrollar uno u otro. ¿Qué será mas interesante: estudiar a los árboles o al bosque? Difícil decidirlo. Los dos enfoques representan también un hueco en la literatura existente en nuestro país, y ambos se percibían como igualmente relevantes para nuestro medio. Por un tiempo seguí investigando ambas vertientes y acumulando notas y datos al respecto.

En un momento dado, decidí que podría intentar desarrollar ambos enfoques en esta obra, y así empezó ésta a organizarse. Por supuesto, ello implicaba profundizar con menos detalle en cada uno de ellos, pero se cuidó siempre que la información tratada fuera siempre la mínima que permitiera en cada capítulo poder hacerse una buena idea del tema, la más relevante y representativa, sin menoscabo, en lo posible, de la calidad y pertinencia de la misma.

De esta forma el trabajo fue dividido en dos grandes partes: la primera cubre entonces ese primer enfoque mencionado, el reduccionista o puntual, donde el objeto de estudio y modelo es una biblioteca digital y en donde se plantean las preguntas enunciadas previamente y, por supuesto, se pretende darles respuesta. Como se mencionó, el lector a quien va dirigida es el profesional o estudiante de bibliotecología que quiere entender la biblioteca digital, aquél que desea conocer cómo se diseña, construye y opera ésta. La presente obra, *Bibliotecas y Publicaciones Digitales* muestra los resultados de la observación de la biblioteca bajo este enfoque.

La segunda parte, abre el horizonte y plantea las cuestiones de la sociedad de la información actual, así como los proyectos y paradigmas que la conforman. Plantea también las preguntas relativas al modelo educativo que tenemos en la actualidad, su evolución, estado del arte y tendencias, y cómo las bibliotecas actuales deben evolucionar para seguir siendo apoyo fundamental a la educación, como lo han sido por mucho tiempo, y para que el puente entre ambas siga sólido y funcional.

Una vez contextualizado el entorno, se procede finalmente a plantear el modelo de organización y funcionalidad de una estructura de bibliotecas digitales a nivel de todo el país, como un ejercicio de planeación a nivel nacional, ante la falta de un modelo a ese nivel al respecto. Los lectores destino por tanto, además de los bibliotecarios, son aquellos relacionados con la planeación en aspectos de sociedad de la información, educación y bibliotecas. Presento este segundo enfoque en la obra *Un Modelo de Planeación de Bibliotecas Digitales Para México*.

Así, pues, el total de la obra puede leerse de varias formas: puede comenzarse por este primer libro si uno está interesado en ese enfoque minimalista y puntual, y si desea conocer la biblioteca digital como un ente, o puede comenzarse leyendo el segundo libro mencionado, si el interés está en los aspectos del contexto social y educativo y en conocer la propuesta del modelo en sí. Un libro puede ser leído sin leer el otro sin que ello ponga en riesgo la comprensión del todo. Obviamente una parte complementa a la otra y nada mejor que estudiar tanto los árboles como el bosque.

No pretendo en modo alguno agotar todos los temas relacionados con la biblioteca digital. El tema es de tal forma extenso que sería imposible afirmar siquiera que se ha tratado medianamente bien. Cada tópico que desarrollaba era una puerta que al abrirse conducía a un corredor lleno de más puertas, y cada una de ellas conducía a su vez a otro corredor en forma infinita, como *La biblioteca de Babel* de Borges. La biblioteca digital es sin duda a su vez una biblioteca de Babel, misteriosa, infinita, complicada, pero por lo mismo interesante y con innumerables secretos aún por descubrir y por explicar.

No pretendo tampoco afirmar que el modelo propuesto en la segunda obra es el mejor o el único modelo posible; no aspiro a que se vuelva un modelo oficial. Mi única intención al respecto es demostrar que mediante un método de planeación adecuado, moderno, coparticipativo e integral es posible diseñar un modelo articulado a nivel de todo el país, con objeto de lograr que nuestras bibliotecas, digitales por lo que se verá, se inserten realmente en un verdadero desarrollo de nuestra sociedad. Si alguien que pueda tomar decisiones al respecto encuentra algo útil en estas obras y lo aplica, por poco que sea, habrán cumplido su propósito y habrá valido la pena escribirlas. Al fin y al cabo a todos nos interesa el desarrollo social, cultural y económico de nuestro país.

Capítulo 1

La biblioteca digital

*“Si no sabes hacia dónde vas, cualquier camino sirve”.
Alicia en el País de las Maravillas,*

de Lewis Carrol. 1865. Capítulo VI.

1.1 CONCEPTO



En el año de 1993 apareció en el mercado un juego para computadora llamado *MYST*, editado por la compañía *Cyan*. Perteneció al tipo de juegos conocidos como *Quest*, o de saga o búsqueda. El juego comienza con una persona que cae en el espacio; al mismo tiempo, el personaje pierde un libro, el cual se desintegra, hoja por hoja, y cae desperdigándose en una isla que existe abajo. El personaje del juego, una vez en la isla, debe recuperar las pá-

ginas del libro para reconstruir la información que contenía. De hecho éste es el mismísimo objetivo del juego, si bien éste nunca se hace explícito. Sucede que el usuario pasa un buen tiempo en el juego antes de percatarse de ello. No hay instrucciones, sólo tres pistas para jugarlo.¹

El juego no es de guerra, no consiste en eliminar extrañas criaturas o evitar trampas, no hay armas ni luchas de artes marciales, y nadie muere. Para jugarlo, uno puede moverse en cualquier dirección en la pantalla, así como entrar y salir en muchos lugares que son, por supuesto, paisajes virtuales de computadora. Se puede tomar una página hallada del libro y leerla. Lo interesante del juego es resolver el misterio en sí mismo: encontrar y reconstruir la información diseminada. El método es búsqueda, análisis, síntesis; inferencia, ensayo y error. El conocimiento se va acumulando con objeto de ir avanzando en la solución

¹ Cyan Worlds. 2001. *Página Web*. Disponible noviembre 3, 2005 en: <http://www.cyan.com>

del misterio. Hay muchos elementos distractores o engañosos. Felizmente, el primer edificio que se encuentra en el juego es una biblioteca, en la cual hay elementos de ayuda, aunque no obvios. Pero la biblioteca es parte importante de la solución.

Si bien el juego es interesante en sí mismo, lo que más me gusta de él es la analogía que plantea con el “mundo real” de la información electrónica. Muchas personas navegan a través de él, pero no siempre saben muy bien en dónde están, hacia dónde van o por qué se encuentran ahí. Es un enigma cuya esencia está contenida en sí mismo. Igual que en el juego, muchos aprenden a navegar en el mundo de la información electrónica por ensayo y error, preguntando e imitando a otros intuitivamente. Como en el juego a veces se olvida que el objeto de navegar es recuperar información necesaria, y no navegar por el sólo hecho de hacerlo y sin saber realmente qué estamos haciendo y cómo llegamos ahí. Del mismo modo, en la vida real de la información electrónica hay elementos distractores, información que es irrelevante para nuestro propósito, y otra que es sesgada o incluso falsa. Información que está ahí, pero que está oculta o su relevancia no es evidente. Tal como en el juego, la biblioteca se encuentra en ese mundo virtual para ayudarnos, pero no son evidentes su presencia ni su utilidad.

Al igual que en el juego, en el “mundo real” de la información electrónica no queda claro dónde está cada pieza de información; ese mundo electrónico permite ver páginas y páginas de información en donde no es evidente cómo buscar, cómo llegamos ahí, o cómo aparece ahí ese material. ¿Toda la información de la red mundial puede considerarse una gran biblioteca? y si no es así ¿dónde comienza y termina la información de las bibliotecas en el ciberespacio? ¿qué tan precisa, seria y profunda es la información fuera de las bibliotecas propiamente dichas?

Curiosamente, hoy día se habla mucho sobre bibliotecas digitales; no sólo por parte de profesionales de la información, también por parte del gran público. *Biblioteca Digital* se ha convertido en la actualidad en un lugar común en el mundo de la información. Y digo curiosamente, porque a pesar de ser uno de los términos más utilizados en nuestro tiempo, no está clara cuál es su definición, su paradigma, sus características y peculiaridades.

Su esencia me recuerda a la energía; casi todos saben de ella en este mundo: la producen, la consumen, la pagan, la estudian, la malgastan, la comentan, protestan en su favor o en su contra, pelean por ella, etcétera; pero pocos, muy pocos, pueden definirla. La prueba está a la vista. Tómese un grupo de libros de física. Se verá que cada uno presenta una definición distinta de energía. Algunos incluso, no dan una definición de energía. Hablan de cómo se manifiesta, los distintos tipos que hay de ella, para qué sirve, quiénes la han estudiado, cómo se transforma, etcétera; pero no presentan una definición de energía. Y no puede decirse que sea un ente “raro”. Todo lo contrario; es de lo más cotidiano, universal y común, y tiene más de doscientos años que se estudia formalmente, sin contar los enfoques esotéricos o cósmicos, los cuales hacen todavía más confusa la idea.

Lo que sucede es que el concepto de “energía” no es sencillo. Es más fácil usar la energía que definirla. Lo mismo sucede con las bibliotecas digitales. Es más fácil decidir producirlas o usarlas que definirlas. El problema es que con la energía es obvio que la producimos o la utilizamos, y con las bibliotecas digitales no. Cuando decidimos crear una biblioteca digital ¿realmente estamos creando una biblioteca digital? Cuando estamos accediendo a información electrónica y pensamos que estamos en una biblioteca digital ¿realmente estamos dentro de una de ellas, o sólo dentro de algo que se le parece? Dicho de otra forma ¿estamos dentro de una biblioteca virtual, o dentro de algo que *sólo virtualmente* semeja una biblioteca?

Para el profesional de la información promedio, es un hecho conocido que existe una Federación de Bibliotecas Digitales en la Unión Americana; una *Iniciativa de Bibliotecas Digitales I* y hasta una *Iniciativa de Bibliotecas Digitales II*, organizaciones que trabajan desde hace ya varios años en el desarrollo de proyectos que se mueven alrededor de la creación de bibliotecas digitales: [DLF, 2001], [DLI1, 1995], [DLI2, 1998] han sido considerados como algunos de los esfuerzos colectivos más grandes hasta el momento en torno a las bibliotecas digitales. En su sitio electrónico, con algo de trabajo por cierto, puede encontrarse una definición de “biblioteca digital”. En la página de la Biblioteca del Congreso de los EUA [LC, 2001] podemos encontrar otra que no es exactamente igual. Tomemos otra página de una biblioteca

digital importante y reconocida; por ejemplo, la de la Universidad de Indiana. El proyecto se denomina *The Indiana University Digital Library Program (IUDLP)*.² Puede recorrerse también el sitio de arriba abajo sin encontrar una definición de qué significa para ellos su “Biblioteca Digital”. Así podríamos continuar con múltiples ejemplos. En la mayoría de sitios “Web” acerca de “bibliotecas digitales” existe escasa o nula alusión sobre qué es una biblioteca digital para ellos, y se asume que todos los usuarios de este planeta lo saben ya. En algunos sitios, no particularmente numerosos, se encuentran algunos documentos, la mayoría no muy recientes, con alguna definición o trazo de lo que significa para ellos la biblioteca digital, los cuales por cierto, no están a la vista de manera evidente o sólo tocan la definición de manera superficial. Deanna Marcum lo concluye así en su revisión de bibliotecas digitales: “...aun cuando muchas bibliotecas están activamente involucradas en proyectos de biblioteca digital, no está muy claro el significado de las mismas”.³

¿Significa esto entonces que el término es demasiado obvio para tener que definirlo detalladamente? Biblioteca es biblioteca y todos saben ya qué es eso, y digital es una representación numérica o con dígitos en forma electrónica y por supuesto ya todos saben también qué significa. Una biblioteca digital es entonces una biblioteca que ofrece documentos y servicios electrónicos y ya está. Parece que caer en la tentación de utilizar ésta o alguna definición parecida es fácil, puesto que por todos lados aparecen bibliotecas y colecciones digitales sin mayor aclaración. ¿O será que el término es muy complejo y mejor lo utilizamos, al igual que la energía, en vez de perder el tiempo en tratar de definirlo?

Desde mi punto de vista, sucede más bien lo segundo. El concepto no es sencillo y es más fácil entonces usarlo que definirlo. No obstante, conviene tratar de hacer una segunda reflexión acerca de ello e intentar definir a esa “biblioteca digital”, o “bibliotecas digitales”, y tratar de

2 The Indiana University Digital Library Program (IUDLP). 2004. Página Web de la Biblioteca. Disponible noviembre 3, 2005 en: <http://www.dlib.indiana.edu/>

3 Marcum, Deanna B. 1997. “Digital Libraries: For Whom? For What?” En: *The Journal of Academic Librarianship*, March 1997, 82.

entender su esencia, sus características, sus elementos, etcétera, para no limitarse simplemente a su explotación sin más conocimiento de causa ya que, al igual que la energía, jugar con ellas sin conocerlas a fondo puede ser riesgoso.

Mucho se ha escrito en los últimos años acerca de los conceptos relativos a la biblioteca digital; de hecho, éste término es la resultante de otros conceptos e ideas anteriores: *biblioteca automatizada*, *electrónica* o *virtual*; *ciberteca*, *biblioteca sin paredes*, *biblioteca de medios* o *mediateca*, *biblioteca compleja*, etcétera. Todos esos términos fueron acuñados a lo largo de los últimos años en un intento de integrar el concepto de la biblioteca tradicional y sus servicios con las nuevas tecnologías informáticas del cómputo y las telecomunicaciones. Si bien analizados en detalle, y tomando en cuenta el tiempo, cada uno de esos términos tiene su muy particular enfoque, contexto, matices y definición, y existen sutiles diferencias entre ellos; no obstante, vistos en conjunto todos tratan de integrar este concepto sinérgico entre computadora, biblioteca y red, en un intento de representar este nuevo ensamble de la información, sus soportes y sus servicios.

El término *biblioteca automatizada* surgió en la década de los sesenta, cuando las bibliotecas integraron las primeras computadoras a su quehacer. Esencialmente aparecía asociado a procesos masivos y relativamente sencillos como las adquisiciones, la elaboración de juegos de tarjetas catalográficas, bibliografías, etcétera. Poco después el concepto *biblioteca electrónica* estaba formado, en términos generales, por objetos físicos que requerían de objetos electrónicos para acceder a aquellos; es decir, se computarizaban los servicios de referencia: índices, resúmenes, directorios, etcétera, los cuales eran consultados a través de *terminales* de computadora; todavía poco o nada los servicios de documentación. La *biblioteca virtual* en esencia era aquella que se organizaba bajo los conceptos establecidos por la “realidad virtual”: una biblioteca que se organizaba en un espacio virtual —*no físico*—, entendiéndose éste como una computadora o conjunto de ellas, de forma distribuida y sin existir un edificio o sede única para las colecciones. No había límite en el espacio para servicios y usuarios.

Detallados estudios se han hecho ya al respecto tratando de establecer esas definiciones, evolución, diferencias, etcétera, por parte de

conocidos autores [Borgman, 1999], [Bawden y Rowlands, 1999], etcétera. Una interesante reseña de esta evolución de los términos y sus significados asociados se encuentra en español en la obra de Araceli Torres, *La Biblioteca Virtual: ¿Qué es y qué promete?* [Torres, 2000]; por lo mismo no haré sino un análisis somero de cada término y sus similitudes o diferencias con el fin de establecer simplemente que estos términos no son sinónimos absolutos, han ido evolucionando hacia un contexto más complejo y han sido reemplazados por el que nos ocupa y que en la actualidad es el más completo y aceptado en el medio. Me referiré a este concepto entonces de ahora en adelante como *biblioteca digital*, ya que en la actualidad engloba el concepto más evolucionado de este tipo de bibliotecas y servicios, y empieza a ser ya del dominio público, más allá de los especialistas de la información, así que será del que me ocuparé con el mayor detalle posible. Esto es sumamente importante, ya que como Borgman apunta:

“las confusiones acerca de la terminología crean enormes obstáculos en la evolución de las tecnologías de información... la terminología usada por los expertos en este campo no tiene por qué ser consistente siempre y cuando cada autor haga explícitas las definiciones que él usa”. [Borgman, 1999].

Biblioteca digital muy frecuentemente se asocia también con una “gran biblioteca total”. Éste es un sueño que muchos han compartido, otros tantos han idealizado y nadie ha visto todavía. Uno de ellos fue Jorge Luis Borges en “La biblioteca de Babel”:

“Cuando se proclamó que la Biblioteca abarcaba todos los libros, la primera impresión fue de extravagante felicidad. Todos los hombres se sintieron señores de un tesoro intacto y secreto. No había problema personal o mundial cuya elocuente solución no existiera...”.⁴

El núcleo de esta visión ha consistido y persistido en una vasta e ideal colección de información que permita el acceso instantáneo al

⁴ Borges, Jorge Luis. 1941 de “El Jardín de los senderos que se bifurcan” en: *Ficciones* (1944). Buenos Aires, Alianza Editorial. Disponible noviembre 3, 2005 en: <http://babel.webcindario.com/>

conocimiento universal. Tal idea no es nada nueva. El mismo Borges nos remite hasta los textos de Aristóteles, particularmente su primer libro acerca de la metafísica; en especial aquel pasaje que expone la cosmogonía de Leucipo: la formación del mundo por la fortuita conjunción de los átomos. Continúa con Demócrito, Pascal, Swift, Carroll, Fechner y Huxley. En el famoso documento llamado “*Carta de Aristeas a Filócrates*” escrito en el siglo II a.C. su autor, Aristeas, atribuye a Demetrio de Falerum, filósofo, discípulo de Aristóteles y “Secretario de Cultura” del rey griego de Egipto Ptolomeo Sóter, el haber fundado un siglo antes la Biblioteca de Alejandría con el propósito de reunir todos los documentos del mundo en una sola colección. Aristeas narra:

“Demetrio de Falerum, presidente de la biblioteca real de Ptolomeo, recibió vastas sumas de dinero con el propósito de reunir, tanto como fuese posible, todos los libros del mundo, ya sea por medio de la compra o de la transcripción. Así lo hizo, en cuanto a él dependía, atendiendo al deseo del rey. En una ocasión le preguntó ¿cuántos libros hay en la biblioteca? —más de doscientos mil, oh, rey, y mis esfuerzos próximos van encaminados a recopilar lo restante, de tal suerte que llegues a tener quinientos mil—. [The Letter of Aristeas..., ca. 150 b.C.]⁵

Como puede verse, la idea de ese acceso universal e inmediato a toda la información posible tiene nada menos que veintitrés siglos de haber sido documentada por primera vez, y ya entonces no era una idea nueva en absoluto: puede remontarse todavía cuatro siglos más atrás a las grandes bibliotecas asirias cuyas enormes colecciones eran legendarias para entonces.

Como este ejemplo pueden trazarse varios a lo largo de la historia antigua y de la edad media. No viene al caso documentar cada uno de ellos, pero vale la pena comentar que en todos persiste ese sueño de

5 “The Letter of Aristeas to Philocrates”. ca. 150 b.C. En: *The Forgotten Books of Eden*. Chapter I, Paragraphs 11-13. Translated in the late 1800’s by Dr. S. C. Malan and Dr. E. Trumpp. Translated into King James English from both the Arabic version and the Ethiopic version which was then published in *The Forgotten Books of Eden* in 1927 by The World Publishing Company. Disponible noviembre 3, 2005 en: <http://www.earth-history.com/Pseudepigrapha/FB-Eden/aristeas-intro.htm>

inclusividad total de la información y acceso cuasi-instantáneo a ella. De hecho, casi todas las culturas que han dependido en buena medida de la palabra escrita han compartido la quimera de la gran biblioteca universal, ésa que poseería el total de la información. Todas ellas se han quedado cortas, y por mucho, de ese propósito de recolectar toda su producción de la palabra escrita, ni siquiera por un breve periodo. Aun las grandes bibliotecas depositarias han contenido sólo una pequeña fracción del material producido por su sociedad [O'Donnell, 1994, p.5]. No obstante, con esa visión han crecido muchas personas e instituciones por milenios tratando de personificar esa gran biblioteca universal, como Borges la ha imaginado, muy a su estilo, en su otro cuento “La Biblioteca Total”⁶:

“Todo estará en sus ciegos volúmenes. Todo: la historia minuciosa del porvenir, Los ‘egipcios’ de Esquilo, el número preciso de veces que las aguas de Ganges han reflejado el vuelo de un halcón, el secreto y verdadero nombre de Roma, la enciclopedia que hubiera edificado Novalis, [...] las paradojas que ideó Berkeley acerca del tiempo y que no publicó, los libros de hierro de Urizen, las prematuras epifanías de Stephen Dedalus que antes de un ciclo de mil años nada querrán decir, el evangelio gnóstico de Basílides, el cantar que cantaron las sirenas, el catálogo fiel de la Biblioteca, la demostración de la falacia de ese catálogo...”.

Cabe mencionar que a su vez, Borges se inspiró para ello en otra historia previa al respecto: *La Biblioteca Universal*, de Kurd Lasswitz, quien en 1901 escribió un cuento donde se calcula matemáticamente el número de volúmenes que ocuparían todos los libros escritos y por escribirse:

“En un momento le calcularé cuántos volúmenes se necesitarían para constituir una Biblioteca Universal [...] —¿Dice Usted que todo estará en esa biblioteca? ¿Las obras completas de Goethe? ¿La Biblia? ¿Las obras de todos los filósofos clásicos? —Sí, y con todas las variaciones en las que nadie ha pensado aún. Encontrarías las obras perdidas de Tácito y su traducción

6 Borges, Jorge Luis. 1939. “La Biblioteca Total”, en: Revista *Sur*. Buenos Aires. No. 59(8), agosto 1939. pp. 13-16. Disponible noviembre 3, 2005 en: <http://www.literatura.org/Borges/LaBibliotecaTotal.html>

a todos los idiomas, vivos y muertos. Además, todas las obras futuras de mi amigo Burkel y mías, todos los discursos ya olvidados, y los que aún deben ser pronunciados, de todos los parlamentos, la versión oficial de la Declaración Universal de la Paz, la historia de todas las guerras subsiguientes, todas las redacciones que todos nosotros escribimos en el colegio y en la universidad—...”.⁷

Se afirma que a su vez Lasswitz se inspiró en el filósofo alemán Gustav T. Fechner, pero ya no abundaré más al respecto.

En los tiempos modernos y más allá de la ficción, considerada —por cierto erróneamente— como la primera referencia a esa biblioteca con acceso universal al conocimiento, pero ya basada fundamentalmente en la tecnología el artículo de Vannevar Bush menciona algo similar en *Como Podríamos Pensar*, (*As We May Think*) escrito a mediados de 1945, casi al final de la segunda guerra mundial [Bush, 1945]. Él era el coordinador de múltiples equipos de científicos que apoyaron el esfuerzo bélico norteamericano durante ese conflicto, y por lo mismo estaba al tanto de los más recientes desarrollos tecnológicos en los laboratorios y centros de investigación del momento. Ya para terminar la guerra decidió hacer un ensayo en función de esos desarrollos acerca de cómo veía él la vida en un futuro tecnificado y, entre muchas otras cosas, concibió una máquina para acceder al conocimiento universal. Él nunca pretendió construir esta máquina; sólo proporcionar una visión de lo que era posible y plausible en función del adelanto tecnológico existente en el momento. No obstante, esta visión es tan asombrosamente parecida a lo que hoy se observa, a pesar de haber sido escrita hace sesenta años, que por lo mismo ese documento es considerado por muchos como la primera referencia a las bibliotecas digitales. Transcribiré algunos párrafos para ilustrarlo. [Las notas en cursivas son comentarios míos]:

Considere un futuro dispositivo para uso personal, el cual es una especie de archivo mecanizado privado y biblioteca a la vez. Necesita un nombre,

7 Lasswitz, Kur d. 1901. *La Biblioteca Universal*. Traducido del original *Die Universalbibliothek*, en: *Traumkristalle*. Varias ediciones. Disponible en español en noviembre 3, 2005 en: <http://www.cinefantastico.com/terroruniversal/ficcion/index.php?t= cuentos&id=45&mode=cuento>

y acuñando uno al azar, se me ocurre llamarlo 'mémex'. Un mémex es entonces un dispositivo en el cual un individuo almacena todos sus libros, registros y comunicados, y está automatizado de tal forma que puede ser consultado con enorme velocidad y flexibilidad. Es una adición enorme e íntima a su propia memoria.

Consiste en un escritorio, y es al mismo tiempo un equipo que puede usarse como una estación de trabajo a la vez que puede ser operado a control remoto. En la parte superior cuenta con pantallas translúcidas, en las cuales el material puede ser desplegado para una conveniente lectura. *(La televisión apenas era conocida por el público entonces, y menos todavía los monitores)*. Hay un teclado, y grupos de botones y palancas; por lo demás su aspecto es el de un escritorio ordinario.

En un extremo está el material almacenado para consulta. El problema de lo voluminoso de la información ha sido solucionado ya que ésta reside en microfilm mejorado. Sólo una parte del interior del mémex está dedicado al almacenamiento; el resto es el mecanismo. Aun si el usuario introdujera cinco mil páginas diarias de materiales le tomaría cientos de años llenar el repositorio, así que puede ser generoso e introducir materiales con liberalidad.

La mayoría de los contenidos del mémex han sido comprados en microfilmes editados previamente y listos para su inserción. Libros de todo tipo, imágenes, revistas, periódicos, etcétera son así obtenidos y puestos en su lugar. La correspondencia de negocios sigue la misma ruta. Y todavía sobra espacio para inserciones propias en forma directa. Sobre el mémex, existe una pantalla transparente sobre la cual pueden ponerse notas, textos, fotografías, etcétera. La acción de un botón hace que sea fotografiado y almacenado en el siguiente espacio disponible en una sección de la película del mémex. *(excelente definición para un no inventado escáner en 1945)*. Se utilizará la técnica de fotografía en seco. *(Así se denominaba en ese entonces a la fotocopia)*.

Se ha previsto, por supuesto, la consulta del registro por el usual sistema de indizado. Si el usuario desea consultar un cierto libro, teclea su código en el teclado, y la cubierta del libro aparece rápidamente ante sus ojos, proyectada en una de sus pantallas. Los códigos frecuentemente usados son nemónicos, así que él no consulta a menudo su libro de códigos. No obstante, cuando lo necesita, un simple teclazo lo trae a la pantalla para su uso. Además, él tiene botones adicionales para esta tarea. Cuando él presiona uno de ellos un poco a la derecha, él pasa las páginas hacia adelante a una gran velocidad, apenas visibles lo suficiente para reconocerlas; un poco más de presión al botón hace que pasen diez páginas a la vez, y más presión hace que pasen cien a la vez. Lo mismo sucede hacia la izquierda, regresando las páginas.

Un botón especial lo transfiere inmediatamente a la primera página del índice. Cualquier libro de su colección puede así ser traído y consultado con mucha mayor facilidad que si fuera tomado de un estante. Como él dispone de varias pantallas, puede ver simultáneamente varias obras (*nótese el uso de ventanas*). Él puede agregar a esas páginas notas al margen, comentarios, tal como si tuviera la página real frente a él, sacando ventaja de algún tipo de fotografía en seco, posiblemente con un accesorio tipo pluma, como el utilizado hoy en día en las estaciones del ferrocarril.

[...] un paso inmediato hacia la indización o archivado de tipo asociativo, cuya idea básica consiste en lograr que cada uno de los elementos seleccione o busque, según nuestra voluntad, a otro elemento de una manera inmediata y automática. Ésta constituye la característica esencial del *mémex*; el proceso de asociar dos o más elementos juntos es lo verdaderamente importante de él. (*Se contemplaba ya el uso de hipervínculos*).

[...] El usuario, con pulsar tan sólo una tecla, hace que los dos elementos queden enlazados de manera permanente. De ahí en adelante, cada vez que el usuario tenga a la vista uno de los elementos, puede llamar al otro en un instante, con sólo oprimir un botón situado bajo el correspondiente espacio del código. Más que eso, una vez que numerosos elementos han sido enlazados entre sí para conformar un sendero de información, pueden consultarse uno tras otro, rápida o lentamente según se desee, simplemente moviendo una palanca similar a la que se usa para pasar las páginas de un libro. Es exactamente igual que si los distintos elementos físicos hubiesen sido colectados a partir de fuentes muy separadas entre sí, y luego encuadrados para conformar un nuevo libro. Y todavía más: cada uno de esos elementos puede pertenecer a su vez, a múltiples senderos de información...

¿No es impresionante la exactitud de la visión con respecto al acceso a la información electrónica en la actualidad? Considérese el contexto en el que fue escrito, en 1945: la única computadora que existía en ese momento, la ENIAC I, no pasaba de ser un ente experimental. Nadie había planteado la posibilidad de manejar información documental en ella; había sido diseñada y utilizada expresamente para hacer cálculos de parábolas de proyectiles; medía treinta metros, pesaba treinta y dos toneladas, tenía casi dieciocho mil bulbos y *no* tenía pantalla alguna. La televisión comercial apenas existía y muy contadas personas tenían TV en sus casas. Empero, Bush concibe ya el despliegue por pantalla a gran velocidad, un escáner, una colección masiva de documentos académicos y notas propias; incluso plantea los hipertextos. Si se sus-

tituye el interior del mémex basado en microfilmes por discos duros, CD-ROM y módem, la visión es de una precisión asombrosa. No es en vano que en las líneas de antecedentes de la red digital mundial el trabajo de Bush es considerado casi siempre como el punto de partida, y por esta razón lo he presentado en primer lugar.

No obstante, en tiempos recientes se ha empezado a difundir una referencia previa tan impresionante como la de Bush, aunque todavía muy poco conocida, probablemente porque su obra magna, la *Clasificación Decimal Universal*, elaborada a partir de la de Dewey, no fue traducida al inglés durante muchos años, y porque no tiene un ensayo puntual y específico de pronóstico tecnológico como el de Bush, sino detalles a todo lo largo de sus obras. En 1934, Paul Otlet publicó su obra magistral *Tratado de Documentación —Traité de Documentation: Le Livre Sur Le Livre—* el cual constituye una síntesis de sus cuarenta años de experiencia respecto a este tema. Otlet había fundado en Bruselas, en 1892, la “Oficina Internacional de Bibliografía” —*Office International de Bibliographie*— la que se convirtió tres años después en el “Instituto Internacional de Bibliografía” —*Institut International de Bibliographie*— y finalmente, en 1931, en el “Instituto Internacional de Documentación” —*Institut International de Documentation*— cuyo último cambio de nombre reflejaba la importancia que iba adquiriendo la información y su explotación. Él fue el primero en comprender el problema que el incremento cada vez mayor de documentos a nivel mundial implicaría en cuanto a su control y recuperación.

En esta obra, Otlet presentó una imagen muy visionaria del futuro de la documentación, tanto como la imagen de Vannevar Bush, pero once años antes que aquél. Visualizó las posibilidades de la tecnología como apoyo a la consulta documental, en particular las de la televisión —cuatro años antes de la primera transmisión comercial—. Él fue el primero en señalar la importancia y la futura multiplicación de los documentos sonoros y el advenimiento, en un futuro, de los documentos táctiles, olfativos y gustativos. Previó las aplicaciones del microfilm como medio de preservación en bibliotecas y definió *la estación de trabajo del investigador*: un ente hipermedios que debería integrar instrumentos de trabajo auxiliares al trabajo intelectual, tales como la transcripción de voz a texto, la teleconferencia, los textos a distancia

con ayuda del teléfono, la “hiperdocumentación” —hoy hipertextos— “...como una extensión de las capacidades del intelecto”. Él identificó ahí los factores que caracterizaron la nueva disciplina de la documentación: la existencia de una parte encargada de estudiar los soportes, y otra centrada en las máquinas que la harían posible. Esta última vertiente recibió el clarificador nombre de *documentación automática* [Otlet, 1934]. Al estudiarlas con cuidado en su contexto de época, las visiones de Otlet resultan tan o más impresionantes que las de Bush.

La siguiente referencia trascendental de lo que se denomina hoy biblioteca digital se realizó aparentemente en los Estados Unidos, en el Instituto Tecnológico de Massachusetts —MIT— a principios de los sesenta, bajo la dirección de Joseph Licklider, cuyos resultados fueron presentados en su libro *Libraries of the future* [Licklider, 1965]. Esta notable obra escrita en 1959 y editada en 1965, la cual por cierto fue dedicada a Vannevar Bush, describe el diseño de lo que él llamó “sistemas precognitivos para el año 2000”. En ella establece:

...tenemos que substituir al libro por un dispositivo que haga fácil transmitir la información sin transportar el material, y que no tan sólo presente la información a las personas sino que también la trate para ellos; para poder proporcionar estos servicios, se requiere claramente de una mezcla entre biblioteca y computadora.

En esta obra se esbozaban ya las publicaciones y las bibliotecas digitales basadas en “esquemas de procesamiento de información” las cuales estarían libres de las restricciones físicas de los libros y las estanterías. Algunos de sus enunciados caían más en el campo de la inteligencia artificial que en del procesamiento de información, pero en ese entonces los algoritmos para imitar ciertos procesos del razonamiento humano estaban en pañales; no obstante, Licklider tenía confianza en que en 35 años evolucionarían lo suficiente para que muchos de los procesos intelectuales de la biblioteca fueran realizados por medio de una máquina. Si bien esto dista mucho de ser una realidad hoy día y el estudio de este autor ha quedado sesgado en este aspecto, otros avances espectaculares en diversas ramas de la electrónica, como ha sido la miniaturización de componentes, la enorme capacidad de los

equipos, sobre todo de los de almacenamiento, y la dramática caída en los precios del equipo han hecho que algunas de sus estimaciones, de manera indirecta, no estén tan lejos de la realidad.

También en 1965 fue enunciado por Ted Nelson un proyecto llamado *Xanadú*, diseñado para ser un *depósito en-línea para todo lo que alguien ha escrito alguna vez*, un *hipertexto realmente universal*. La idea de hipertexto es esencialmente *texto de escritura no secuencial que se bifurca sucesivamente y permite opciones al lector, de preferencia sobre una pantalla interactiva*. [Nelson, 1974]. Él también atribuye la inspiración del concepto del hipertexto a Vannevar Bush, afirmando en uno de sus escritos Bush tenía razón al respecto. Él confiaba en que un día la ciencia desarrollaría dispositivos de almacenamiento capaces de contener todas las obras publicadas sobre la tierra. Un rasgo interesante de *Xanadú* es que un documento, una vez insertado en el sistema, nunca sería suprimido; incluso si una nueva versión de un documento específico fuera insertada, la vieja versión sería guardada porque podría haber eslabones que lo señalaron. Este rasgo de *Xanadú* podría ser usado para el direccionamiento de versiones diferentes —por ejemplo para el desarrollo de programas de computadora—. Al final, casi cuarenta años después, lo único práctico logrado por Ted Nelson y su proyecto fue acuñar el término *hipertexto*. El problema es que Nelson no lo expresó como una visión, al igual que sus antecesores, sino que dio la impresión al mundo de que ya estaba construyendo el sistema y que en poco tiempo entraría en servicio; por ello en la actualidad es considerado como una de las grandes utopías publicitarias de la industria del cómputo y la información y por lo mismo su proyecto ha sido fuertemente debatido y demeritado con el tiempo. En la realidad, el primer hipertexto fue el sistema “NLS” —*ON Line System*— creado en 1968 en el Instituto de Investigaciones de Stanford —*Stanford Research Institute*— por Douglas Engelbart, inventor de sistemas de texto electrónicos y del ratón para computadora. Una reseña histórica mucho más completa acerca de los hipertextos puede encontrarse en el excelente trabajo de Guy Teasdale *L'Hypertexte, Historique et Applications en Bibliothéconomie —El Hipertexto, Historia y Aplicaciones en Biblioteconomía—*. [Teasdale, 1995].

*“Si, (como el griego afirma en el Crátilo)
El nombre es arquetipo de la cosa,
En las letras de rosa está la rosa
Y todo el Nilo en la palabra Nilo”.
“El Golem”, en El otro, el mismo,*

Jorge Luis Borges

1.2 DEFINICIÓN

Como hemos establecido, la computadora fue entrando cada vez más a lo largo de estas últimas cuatro décadas al ámbito de las bibliotecas en todo el mundo, y conformando nuevos conceptos alrededor de la fusión de estos dos ambientes. En un principio era sólo la impresión de tarjetas catalográficas. A continuación fueron catálogos electrónicos que en pocos años pudieron ser consultados en línea a través de “terminales” de grandes sistemas centrales de biblioteca. Con el advenimiento de la computadora personal en estas dos últimas décadas el acceso a más y más información bibliográfica se fue acrecentando. Desde que el connotado profesor de Harvard y director del *Harvard Business Review*, Theodor Levitt, acuñara el término *globalización* en 1983, se le daba nombre a una nueva configuración destinada a modificar profundamente los diversos ámbitos en todo el orbe. Originalmente, la idea de Levitt se orientaba sólo a la globalización de los mercados; esto es, a operar en el mundo con una nueva lógica mercantil, y se circunscribía a un nuevo modelo tecnoeconómico que suponía la reestructuración del capitalismo a nivel del planeta. Dos décadas después, y con el advenimiento del Internet, la *Web* y las telecomunicaciones, es claro que nos vemos enfrentados, además, a una profunda mutación política y cultural. Hablar entonces de globalización supone entonces explorar un nuevo modelo social y al mismo tiempo un modelo inédito en el desarrollo mundial. Me referiré a ello con detalle posteriormente.

Estos elementos, y algunos más, fueron conformando ese concepto de fusión entre la biblioteca tradicional, sus colecciones y servicios, con los equipos electrónicos y el teleproceso de datos, y consolidando definitivamente la sinergia de estos elementos bajo un nuevo modelo

de desarrollo. Como ya hemos establecido, al tratar de definir las relaciones y condiciones que se crean con esta fusión se fueron acuñando con el tiempo varios neologismos que han tratado de definirla. Así, vimos nacer los términos *biblioteca electrónica* o *virtual*, *biblioteca sin paredes*, el muy europeo *ciberteca*, las *mediatecas* o *bibliotecas de medios*, *biblioteca compleja*, etcétera, los cuales, como ya hemos mencionado, fueron desembocando hacia el término más general, moderno y aceptado de *biblioteca digital*, y el que por tanto trataré de definir para referirme a este tipo de bibliotecas. Un estudio de [Watstein, Calarco, Ghapery, 1999] acerca del uso de los términos, demuestra que el concepto de biblioteca digital presenta una casi nula ocurrencia durante los setenta, un pequeño incremento en los ochenta y una explosión desde 1994, lo cual lo posiciona en su evolución y lo relaciona con el advenimiento de la red *www*, en donde todas las bibliotecas digitales se encuentran.

Con respecto a los primeros términos, el *Diccionario en-Línea de Bibliotecología y Ciencias de la Información* ODLIS —*On-Line Dictionary of Library and Information Science*— de la Universidad Estatal Occidental de Connecticut establece que:

una biblioteca sin paredes es aquella en la que las colecciones no existen sobre papel, microforma, ni en ninguna forma tangible, pero es electrónicamente accesible en formatos digitales. Tales bibliotecas existen sólo a una escala muy limitada, pero en las bibliotecas tradicionales de los Estados Unidos basadas en impresos, los catálogos y los índices y resúmenes están disponibles en-línea; algunas revistas y obras referenciales en texto completo pueden estar también disponibles electrónicamente. Algunas bibliotecas y sistemas bibliotecarios se autonombren ‘virtuales’ por el simple hecho de ofrecer servicios en-línea. El término de ‘biblioteca digital’ es más apropiado ya que el término ‘virtual’ —proveniente del concepto de “realidad virtual”— sugiere que la experiencia de usar tal biblioteca no es la misma que en una ‘verdadera’. Si bien la experiencia de lectura o inspección de un documento sobre una pantalla de computadora puede ser cualitativamente diferente de leer la misma publicación en forma impresa, el contenido de información es el mismo independientemente del formato. [ODLIS, 2002]

Roy Tennant establece la diferencia afirmando que:

una biblioteca electrónica es una biblioteca que cuenta con materiales y servicios electrónicos. Tales materiales pueden incluir tanto todos los materiales digitales como una variedad de formatos analógicos que funcionan con energía eléctrica, como por ejemplo cintas de video. Por lo tanto, el término ‘biblioteca electrónica’ abarca todo el material que puede ser guardado por una ‘biblioteca digital’ más otras cosas, y es por tanto más completo. No obstante, está pasado de moda. Una biblioteca digital es una biblioteca que cuenta con materiales y servicios digitales. Tales materiales son documentos almacenados, procesados y distribuidos por equipos y redes digitales (binarios). Los servicios digitales son servicios, por ejemplo el de consulta, que son proporcionados digitalmente sobre redes de computadora. Por lo tanto, ambas bibliotecas, la digital y la electrónica son bibliotecas virtuales sólo si existen virtualmente; esto es, que esa biblioteca no existe en ‘el mundo real’, como por ejemplo, las bibliotecas formadas por la unión de colecciones de varias bibliotecas distintas organizadas en un espacio virtual por medio de computadoras y redes. [Tennant, 1999].

¿Qué es lo que realmente significa el concepto de *biblioteca digital*? Como ya hemos mencionado, es un concepto complejo; de hecho, varía sensiblemente según el enfoque disciplinario. Comienza con la percepción que tiene el público en general: el acceso a la información ilimitada con sólo unos golpes a un teclado. Enormes repositorios ominoscentes que pueden contestar prácticamente cualquier pregunta; una especie de “bola de cristal” en donde la esfera ha sido sustituida por una pantalla de cristal —o ahora de plasma— gracias a los avances tecnológicos, concepto que como ya hemos mencionado es muy romántico pero está muy alejado de la realidad.

El problema de definir el concepto comienza con que *biblioteca digital* no significa lo mismo para distintas comunidades no bibliotecarias. Los que [Nürnberg, *et al*, 1995] hacen una descripción de varios enfoques de biblioteca digital que otras comunidades tienen alrededor de esta idea, de entre los cuales podemos destacar los siguientes ejemplos:

- Desde el punto de vista computacional y de la recuperación de información, es una gran base de datos.
- Para las personas que trabajan en tecnología de hipervínculos, es una aplicación particular de los métodos de hipertextos.

- Para aquellos que trabajan en entrega de información en redes de área expandida, es una aplicación de la *web*.
- Para los programadores y desarrolladores de aplicaciones de cómputo, son colecciones de algoritmos computacionales o piezas de programación, reunidos en un solo sitio.
- Para vendedores de bancos de información, revistas y documentos electrónicos sus productos son bibliotecas digitales.

Algunos de los conceptos mencionados en el párrafo anterior podrían ser bibliotecas digitales; otros no. Pero ilustran la idea de que el término puede significar distintas cosas para diversas comunidades. Un ejemplo sumamente espectacular de lo que es considerado una biblioteca digital es el de la “*World Wide Web*” o red mundial. Como la red está acumulando documentos, muchos tienden a llamar a esta gran colección documental “biblioteca digital”, ya que ahí puede hallarse información. De acuerdo con Clifford Lynch,⁸ reconocido académico al efecto, la red no puede ser en su conjunto una biblioteca digital, ya que como él afirma:

A veces escuchamos decir que la Internet es la biblioteca mundial de la era digital. Esta afirmación no se sostiene más allá de una afirmación casual. La Internet, y en particular su colección de recursos documentales conocidos como la World Wide Web, no fueron diseñados para sostener la publicación organizada ni la recuperación de información como lo fueron las bibliotecas. Ha ido evolucionando como un caótico repositorio de la ‘publicación electrónica’ mundial, pero no es una biblioteca digital. Este almacén de información no contiene sólo libros y revistas, sino menús de restaurante, datos crudos, minutas de reuniones, publicidad, transcripción de conversaciones, etcétera...

Michael Gorman⁹ es todavía más directo cuando afirma:

8 Lynch, Clifford. 1997. “Searching the Internet”, en: Scientific American. March 1997. Disponible noviembre 3, 2005 en: <http://www.hackvan.com/pub/stig/articles/trusted-systems/0397lynch.html>

9 Gorman, Michael. 1998. “*What is the Internet*”, en: *The One Person Library*, vol. 15, No. 6 (6/98) p. 5.

...Tome un libro, quítele las cubiertas, quítele la página de portada, quite la tabla de contenido, quite el índice, corte la encuadernación separando todas las hojas y distribúyalas aleatoriamente en una habitación. Ahora trate de hallar ahí la información que necesita... eso es la Internet.

El mismo Gorman abunda en ello al narrar su fábula:

Hay un universo alternativo en el que existen libros pero no documentos electrónicos. En este universo los bibliotecarios no tienen ningún control sobre los libros que compran, no tienen criterios para la selección, ni para los planes de adquisición del material ni para el desarrollo de la colección. En cambio, las veinticuatro horas del día varios camiones van y vienen a la biblioteca y depositan montones de libros no deseados ni solicitados, la mayoría de editores desconocidos, editoriales de revistas del corazón y auto-editores que trabajan en sótanos. Ciertamente, algunos de estos libros podrían ser de interés, pero ¿cuáles, cómo los pueden encontrar los bibliotecarios y los usuarios de la biblioteca y qué hacer con el resto? En este universo diferente, la bibliotecología se convierte en un proceso mucho más aleatorio y desorganizado que cualquier otro. La biblioteca envía cuadrillas de personal entrenado para rebuscar entre los montones intentando encontrar documentos que valga la pena catalogar y archivar. Pero, ¡espere! ¡Es un universo alternativo y, tras haber seleccionado cien libros de los montones y haberlos catalogado y organizado, los bibliotecarios regresan al día siguiente y descubren que veinticinco de ellos han desaparecido y veinticinco tienen otro título! Mientras tanto, los montones que están fuera de la biblioteca se multiplican y cambian de forma; por cada cien libros que el equipo de la biblioteca rescata, los camiones de reparto han añadido doscientos [...] Hay demasiados documentos, algunos desaparecen después de ser registrados, otros cambian sus características, otros son falsos en cuanto que no son lo que dicen ser, algunos no pueden encontrarse y no hay un filtro del material efímero y del falso, como lo hace de alguna forma la industria de la edición y venta de libros. [Gorman, 2001].

Puede observarse también lo que resulta cuando otro tipo de comunidades tratan de definir, bajo su muy peculiar perspectiva, una biblioteca digital propiamente dicha. Como ejemplo, Henry Gladney¹⁰ propone esta definición al término, sin duda muy “computacional”:

10 Gladney, Henry M., *et al.* 1994. *Digital Library: Gross Structure and Requirements*, en: *Proceedings of the First Annual Conference on the theory and practice of Digital Libraries*.

Una biblioteca digital es un ensamble de cómputo digital, mecanismos de comunicación y programación necesarios para reproducir, emular y extender los servicios provistos por bibliotecas convencionales, como son la catalogación, búsqueda y disseminación de la información.

Hay que contextualizar la definición el año 1994- y a la comunidad que lo emite. Para la comunidad bibliotecaria tales conceptos no corresponden a lo que puede considerarse una biblioteca digital; pero ilustran y demuestran el hecho de que este concepto es amplio y tiene diversos enfoques dependiendo del contexto de la comunidad que lo aplica.

Entrando a los conceptos y definiciones elaborados por comunidades más próximas a las bibliotecas pueden distinguirse diferencias y coincidencias. Puede observarse que algunas de las cuestiones más notables giran alrededor de subrayar el uso de la tecnología; otras, resaltan el desarrollo de colecciones; otras más, el de los contenidos. Unas definiciones son muy superficiales; otras bastante más complejas.

Como ejemplo de las primeras, tenemos versiones que soslayan el punto, o lo tocan sólo por encima. Mintzer comienza un artículo sobre el desarrollo de bibliotecas digitales diciendo:

El término biblioteca digital no está todavía bien definido ni en la literatura popular ni en la técnica, pero generalmente se usa para describir sistemas que manejan colecciones de datos muy grandes y proveen abundantes herramientas de búsqueda y otros servicios al estilo de las bibliotecas.¹¹

Jesús Tramullas indica que, de acuerdo con quienes han tratado de definir el término, pueden distinguirse tres aproximaciones al efecto, cuyos ejemplos iremos analizando:

- Definiciones por características , como la de la *“Association of Research Libraries”*.
- Definiciones por arquitectura o componentes , como la de Borgman y la de UCLA/NSF.
- Definiciones por Objetivos o Funciones , como la de la DLF.

11 Mintzer, Fred. 1999. “Long Term Storage of Digital Media Objects for Digital Libraries”, en: *NSF Workshop on Data Archival and Information Preservation*, Washington DC, March 1999.

Comencemos por las más sencillas; el diccionario ODLIS establece la *biblioteca digital* como:

una biblioteca en la cual una proporción significativa de los recursos está disponible en formatos explotables por computadora, en vez de impresos o en microformatos. En las bibliotecas, el proceso de digitalización comenzó con el catálogo, siguió con los servicios de resúmenes y bibliografías, luego con las publicaciones periódicas secundarias, continuó con las revistas en texto completo y después con la publicación de libros. [ODLIS, 2002]

La Biblioteca Británica en su “Programa de Biblioteca Digital” nos dice que consiste en

el uso de las tecnologías digitales para adquirir, almacenar, conservar y facilitar el acceso a la información y a publicaciones editadas originalmente en forma digital o a documentos existentes en formatos tradicionales, como impresos o audiovisuales, que han sido convertidos a formatos digitales.¹²

Continuemos con una de las más conocidas y, por lo mismo, aceptadas en general, por provenir de quien la emite; la establecida por la “Federación de Bibliotecas Digitales” —“*Digital Libraries Federation*” o DLF— en la unión americana:

Las bibliotecas digitales son organizaciones que proveen los recursos, incluyendo personal especializado, para seleccionar, estructurar, ofrecer acceso intelectual, interpretar, distribuir, preservar la integridad y asegurar la persistencia a lo largo del tiempo de colecciones de obras digitales, de tal forma que estén oportuna y económicamente disponibles para su uso por parte de una comunidad definida o conjunto de comunidades. [DLF, 2001]

El problema con esta definición es que pudo haber sido adecuada para el año en que fue acuñada, 1996, pero no ha sido actualizada y ello crea confusión entre los nuevos interesados y no da buena imagen de la DLF. Nótese que si de esta definición se quitase la palabra “digitales”, que aparece dos veces, la definición sería *exactamente* la de una biblioteca tradicional. Por consiguiente, uno podría estar fuertemente

12 British Library. Página Web de la Biblioteca. Disponible noviembre 3, 2005 en: <http://www.bl.uk>

tentado a suponer que la biblioteca digital es una biblioteca tradicional, pero agregándole el ingrediente de lo “digital”, lo que sea que esto signifique. Veremos más adelante que el concepto de la biblioteca digital no es tan simple como eso y por lo mismo descarto esta definición por superficial.

Otras definiciones revisadas incluyen las de Christine Borgman, la de William Birmingham, la de Gary Marchionini y Hermann Mauer, la de Gary Cleveland, la de Gobinda Chowdhury y la de Peter Graham: como se ha mencionado estas definiciones tienen elementos comunes y disímbolos; destaco algunas de entre ellas:

Para William Birmingham,¹³ el término biblioteca digital:

es un nombre genérico para estructuras dinámicas federadas que proporcionan acceso intelectual y físico a redes crecientes de información codificada en formatos digitales multimedios.

Y según él mismo, las características que distinguen las bibliotecas digitales de las convencionales son las siguientes:

- Proporcionan información en cualquier tiempo y lugar.
- Proporcionan acceso a colecciones de información multimedios, construidas gracias a la integración de texto, imágenes fijas, gráficas, sonido e imágenes en movimiento.
- Permiten el ingreso y la representación personalizada de la información, incluyendo el soporte para recuperar sólo información relevante.
- Mejoran las actividades colaborativas tales como la investigación, el aprendizaje y diseño, al reducir las barreras del tiempo y la distancia.

Gary Cleveland nos da también una caracterización de lo que son las Biblioteca Digitales.¹⁴

13 Birmingham, William P. 1995. An agent-based architecture for digital libraries, en: *D-Lib Magazine*. (Junio 1995.)

14 Cleveland, Gary. 1998. Digital Libraries: Definitions, Issues and Challenges. (UDT Occasional Paper # 8). *IFLANET: Activities & Services*. Página Web Oficial de la IFLA. Disponible noviembre 3, 2005 en: <http://www.ifla.org/VI/5/op/udtop8/udtop8.htm>

- Las bibliotecas digitales son la cara digital de las bibliotecas tradicionales que incluyen tanto colecciones en soportes tradicionales como colecciones digitales, así que abarcan materiales electrónicos y en papel.
- Las bibliotecas digitales incluirán también los materiales digitales que existen fuera de los límites físicos y administrativos de esa biblioteca.
- Las bibliotecas digitales incluirán todos los procesos y los servicios que son actualmente la columna vertebral de cualquier biblioteca. No obstante, dichos procesos tradicionales, aunque forman la base del quehacer bibliotecario en el ámbito digital, tendrán que ser revisados y optimizados para conciliar las diferencias entre los nuevos medios digitales y los soportes fijos tradicionales.
- Las bibliotecas digitales proporcionan idealmente una vista coherente de toda la información contenida dentro de una biblioteca, sin importar su soporte ni formato.
- Las bibliotecas digitales servirán a comunidades o grupos particulares, tal como las bibliotecas tradicionales lo hacen ahora, si bien esas comunidades pueden estar ampliamente dispersas a través de la red.
- Las bibliotecas digitales requerirán tanto las habilidades de profesionales de la biblioteca como de profesionales del cómputo para ser viables.

Continuando con la caracterización de Cleveland nos dice:

Algo que las bibliotecas digitales nunca serán, es un sistema único, completamente digital que provea acceso instantáneo a toda información para todos los sectores de la sociedad de cualquier parte del mundo; ello es sencillamente irreal. Tal concepto proviene de los días en que la gente no estaba muy al tanto de la verdadera complejidad de construir bibliotecas digitales. En vez de ello, es más probable que serán un conjunto de recursos de diversa índole y variados sistemas, uniendo a comunidades y grupos específicos, creados para fines específicos. También incluirán, quizás indefinidamente, colecciones en papel. Más allá, la interoperabilidad a lo largo de diversas arquitecturas, metadatos y formatos será posible de manera parcial y sólo dentro de sistemas aislados y definidos para propósitos y comunidades preestablecidas.

Es importante detenernos en el primer punto de la caracterización de Cleveland: *“Las bibliotecas digitales son la cara digital de las bibliotecas tradicionales que incluyen tanto colecciones en soportes tradicionales como colecciones digitales. Así que ellas abarcan materiales electrónicos y en papel”*. Bawden y Rowlands lo resaltan mucho en sus estudios e inclusive presentan el término *biblioteca híbrida* ya

que esta situación se presenta en la inmensa mayoría de las bibliotecas actuales. De acuerdo con este concepto la biblioteca digital es:

un espacio físico o virtual, o una combinación de ambos, en la cual una porción significativa de los recursos ofrecidos a los usuarios existe sólo en forma digital [...] la biblioteca provee servicios y materiales en parte físicos y en parte virtuales. Por tanto asume un ideal de integrar y proveer todos sus recursos de información al margen de formatos o medios. [Bawden y Rowlands, 1999, p. 11-13]. Este concepto está muy arraigado en las bibliotecas digitales europeas.

Siguiendo con las caracterizaciones, la “*Asociación de Bibliotecas de Investigación*” de la unión americana —*Association of Research Libraries*—, en su documento *Realizing Digital Libraries* de 1995, no da definición, a pesar de que el documento tiene un apéndice que comienza con el título “Definición y propósitos de una biblioteca digital”; simplemente subraya los aspectos que han sido detectados como comunes en otras definiciones [*Association of Research...*, 1995]:

- La biblioteca digital no es una entidad única.
- La biblioteca digital requiere tecnología para integrar los recursos de muchas de ellas.
- Las ligas entre las muchas bibliotecas digitales y servicios de información deben ser transparentes al usuario final.
- La meta del acceso universal a múltiples colecciones persiste.
- Las colecciones de las bibliotecas digitales no están limitadas a elementos documentales tradicionales; se extienden a elementos digitales que no pueden ser representados en formatos impresos.

Seamans y Mcmillan¹⁵ nos dicen, más como una reflexión-caracterización que como una definición:

Una biblioteca digital debe ser una libre extensión de la biblioteca tradicional proveyendo a sus lectores acceso a información en todo formato

15 Seamans, Nan y Mcmillan, Gail. 1998. *Digital Library Definition for DLI2*. Disponible noviembre 3, 2005 en: <http://scholar.lib.vt.edu/DLI2/defineDL.html>

la que ha sido evaluada, organizada, archivada y preservada. El acceso a este creciente cúmulo de información se da a través de sistemas electrónicos personalizados y servicios de profesionales de la información. La biblioteca digital provee valor agregado y ahorra tiempo reduciendo significativamente el tiempo de acceso. Reduce la necesidad de la proximidad física de los recursos de información, pero hace énfasis en la calidad de esos recursos.

Mel Collier,¹⁶ en el “Simposio Internacional de Investigación, Desarrollo y Práctica en Bibliotecas Digitales” efectuado en 1997, nos define que

una biblioteca digital es un ambiente administrado de materiales multimedios en forma digital, diseñados para el beneficio de su comunidad de usuarios, estructurado para facilitar el acceso a sus contenidos, y equipado con ayudas para navegación en la red global [...] con usuarios y acervos totalmente distribuidos, pero administrados como un todo coherente.

Jesús Tramullas,¹⁷ en su curso “Las Bibliotecas Como Sistemas Multifacetados de Información” establece una definición muy sencilla pero muy completa al definir las como

un sistema de tratamiento técnico, acceso y transferencia de información digital, estructurado *alrededor del ciclo de vida de una colección de documentos digitales, sobre los cuales se ofrecen servicios interactivos de valor añadido para el usuario final.*

En la *Conferencia Anual de IFLA* del año 2000, Sharon y Frank¹⁸ definen a la biblioteca digital como *una biblioteca computarizada en la cual la mayor parte de la información es digital*, y además distinguen tres tipos de ellas: 1) la *biblioteca digital unitaria*, que es aquella

16 Collier, Mel. 1997. “Towards a General Theory of the Digital Library”, en: *Proceedings of the International Symposium on Research, Development and Practice in Digital Libraries*: ISDL'97.

17 Tramullas, Jesús. 2002. “Las Bibliotecas Como Sistemas Multifacetados de Información”. Sesión en el programa de doctorado *Informática y Modelización Matemática*, Universidad Rey Juan Carlos, España. Disponible noviembre 3, 2005 en: <http://www.tramullas.com/presentaciones.html>

18 Sharon, Taly y Frank, Ariel J. 2000. Digital Libraries on the Internet, en: 66th *IFLA Council and General Conference*. Jerusalem, 13-18 agosto 2000. Disponible noviembre 3, 2005 en: <http://www.ifla.org/IV/ifla66/papers/029-142e.htm>

biblioteca típica reconvertida a una versión altamente computarizada; 2) la *biblioteca digital federada*, consistente en una federación de varias bibliotecas independientes centradas en un tema común y dentro de una red; 3) la *biblioteca digital cosechada*, que es aquella biblioteca virtual que provee acceso integrado a material desperdigado en la red.

En el Taller del Instituto para Aplicaciones de Tecnologías de Información (IITA) de mayo de 1995, Lynch y García Molina¹⁹ establecieron:

Las Bibliotecas Digitales son vistas como sistemas que proveen acceso coherente a una comunidad de usuarios a vastos y organizados repositorios de información... la capacidad del usuario para acceder, reorganizar y aprovechar este repositorio se ve optimizada por las posibilidades de la tecnología digital...

Otra de las definiciones-caracterizaciones muy completa es la del Grupo de Trabajo D-Lib en Métrica de Bibliotecas Digitales —*D-Lib Working Group on Digital Library Metrics*— pues hace una definición sencilla con seis elementos y a su vez, posteriormente, hace una definición detallada de cada uno de los elementos que la componen para establecerla con toda claridad; separo los seis elementos con diagonales para identificarlos más fácilmente. La definición establece que una biblioteca digital es:

la colección de servicios / y la colección de objetos de información / que apoyan a los usuarios en el manejo de esos objetos / así como la organización y presentación de esos objetos / disponibles directa o indirectamente / por medios electrónicos o digitales.

Los detalles de cada uno de los elementos pueden consultarse en su documento al respecto *The Scope of the Digital Library* [Leiner, 1998].

En el Taller de UCLA-NSF llevado a cabo en 1996 [UCLA-NSF, 1996] se elaboró una definición-caracterización mucho más profunda y completa de las bibliotecas digitales. Su reporte final define:

19 Lynch, Clifford y García-Molina, Héctor. 1995. *Interoperability, Scaling, and the Digital Libraries Research Agenda: A report on the May 18-19, 1995 IITA Digital Libraries Workshop*. Disponible noviembre 3, 2005 en: <http://diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.html#2>.

Las bibliotecas digitales son un conjunto de recursos electrónicos y capacidades técnicas asociadas para la creación, búsqueda y uso de la información. En este sentido son una extensión y mejora de los sistemas de almacenamiento y recuperación que manipulen datos en cualquier medio —texto, sonido, imagen estática o dinámica— existentes en redes distribuidas. El contenido de las bibliotecas digitales incluye datos, metadatos descriptivos y metadatos de hipervínculo. Las bibliotecas digitales están construidas —colectadas y organizadas— por una comunidad de usuarios, y las capacidades funcionales de la biblioteca apoyan las necesidades de información y costumbres de esa comunidad. Son componente de esa comunidad en la cual grupos e individuos interactúan unos con otros, utilizando recursos y sistemas de datos, información y conocimiento. En este sentido ellas son una extensión, mejora e integración de una variedad de instituciones de información en lugares físicos donde los recursos son seleccionados, colectados, organizados, preservados y accesados en apoyo a comunidades específicas de usuarios. Estas instituciones de información incluyen, entre otras, a bibliotecas tradicionales, museos, archivos, y escuelas, extendiendo sus servicios a estos lugares, y además a aulas, oficinas, laboratorios, hogares y áreas públicas.

El grupo que elaboró esta definición fue coordinado por Christine Borgman, de UCLA, quien usa por lo mismo esta definición en sus estudios acerca del tema y es una de las más aceptadas a la fecha por los especialistas[Borgman, 1999]. Como puede observarse, ésta es una super-definición, bastante compleja, y trata de cubrir múltiples aspectos desde muchos enfoques. De hecho es más una caracterización de una biblioteca digital que una definición, y se antoja demasiado elaborada para ser utilizada en la práctica y dirigida a un amplio público. Sin embargo, para una comunidad especializada en bibliotecología u otras ramas afines de la información es una de las más completas que se hayan elaborado y debe observarse con cuidado, y si bien podrían agregarse o precisarse algunos elementos, puede ser utilizada como base para un análisis más profundo de lo que el término biblioteca digital conlleva, y por ello la tomo como punto de partida para los análisis subsiguientes.

En 1997, al final del “Taller de Planeación de Santa Fe Acerca de Medio Ambientes de Trabajo de Conocimiento Distribuido” [Santa Fe Workshop on Distributed Knowledge Work Environments, 1997] se estableció el consenso, cada vez más aceptado a nivel mundial donde:

el concepto de biblioteca digital no es simplemente el equivalente de colecciones digitalizadas con herramientas de manejo de información. Es más bien un ambiente digital para integrar colecciones, servicios y personas en apoyo a un ciclo vital de creación, disseminación, uso y preservación de datos, información y conocimiento. Las oportunidades y retos que motiven ulteriores investigaciones acerca de las bibliotecas digitales deben asociarse a esta amplia visión del ambiente que representan estas bibliotecas.

Estas dos últimas definiciones, además de la de Tramullas, son las que más me satisfacen, la primera por completa; es sin duda una definición adecuada para profesionales del medio. La segunda, porque siendo simple y concisa, rescata su esencia, y si pusiera un poco más de énfasis en la parte tecnológica o digital, podría ser excelente para los no profesionales. No obstante, la reseña ha sido compilada para poder observar la enorme variedad de definiciones y caracterizaciones que existen alrededor del término.

Finalmente, he aquí una lista de los elementos que de acuerdo a múltiples autores parecen ser el consenso a lo largo de muchas caracterizaciones que definen a una biblioteca digital:

- Proveen acceso rápido y eficiente a través de una buena interfaz.
- Pertenecen a una organización estructurada y lógica.
- Apoyan fuertemente la enseñanza y no sólo el acceso documental.
- Sirven a una comunidad o grupo bien definido.
- Unen lo mejor de los recursos tecnológicos y humanos profesionales.
- Proveen acceso gratuito o barato, aunque sea a una comunidad específica.
- Poseen y adquieren una buena cantidad de recursos documentales.
- Trabajan en forma federada o colaborativa con otras bibliotecas.
- Invierten en el desarrollo de sus colecciones.
- Muchas de ellas combinan los servicios a través de colecciones digitales y en soporte tradicional.
- Sus colecciones:
 - Están bien definidas en cuanto a políticas de selección.
 - Son vastas, y perduran a lo largo del tiempo.

- Están bien organizadas documentalmente y bien administradas.
- Contienen variados formatos multimedios.
- Contienen objetos documentales difícilmente obtenibles en otros lados.
- Contienen una vasta colección de documentos digitales de origen.
- Se complementan con las de otras bibliotecas digitales.

1.3 PROBLEMÁTICA

Otro aspecto que es necesario enfatizar en lo tocante a las definiciones, es que están basadas en las características de funcionamiento, tecnología o construcción de las bibliotecas digitales, pero pocas incluyen aspectos del deber ser de las bibliotecas digitales en aspectos tales como los filosóficos, sociales, pedagógicos y políticos; esto es, la visión humanista de la biblioteca digital en un futuro. Hay que hacernos preguntas tales como ¿quiénes y cuántos tendrán acceso a las bibliotecas digitales? ¿en qué forma deben insertarse las bibliotecas digitales en los modelos educativos del futuro? ¿cómo deben verse las bibliotecas digitales en las perspectivas de un desarrollo sostenido de las sociedades a las que sirven? ¿cómo inciden las bibliotecas digitales en un mundo que tiene la tendencia a globalizarse y conformarse en “bloques”? y ¿cómo deben ayudar a mantener los valores y culturas propios de las comunidades que las crean? En suma: ¿qué papel *social* deben jugar las bibliotecas digitales en un mundo que presenta profundos cambios sociales? Algunas de estas respuestas tienen características que deben ser integradas a las definiciones de la biblioteca digital. De otra forma, corremos el riesgo de definir qué son, pero omitir el para qué las queremos.

Si uno de los sinónimos de biblioteca digital es biblioteca del futuro, entonces tenemos también que definir cuál es ese futuro que queremos, al mismo tiempo que definimos a esas bibliotecas para propiciar de alguna manera que los dos elementos converjan en algún punto del espacio y del tiempo.

Amanda Spink presenta algunas reflexiones en este sentido, que nos hacen recordar que en un futuro cercano se prevé un escenario mundial de escasez de recursos naturales, conflictos sociales, crecimiento de la población mundial, migraciones y degradación ambiental, entre otras cosas, y realiza algunas comparaciones entre puntos de vista neo-liberales y humanistas, por demás interesantes. Sin duda, muchas de estas reflexiones deben incluirse en la definición de las bibliotecas digitales del futuro cercano [Spink, 2001].

De lo anterior, puede deducirse que *biblioteca digital* es un ente complejo, que conlleva *insoslayablemente* un concepto social que va mucho más allá de la simple computarización de los acervos y servicios de una biblioteca tradicional, como muchos todavía piensan. Esto es explicable: puede observarse que las sociedades tratan de encuadrar por inercia los nuevos conceptos derivados de los adelantos tecnológicos de acuerdo con los enfoques tradicionales de la época que conocen.

Así cuando el cinematógrafo fue descubierto durante un buen número de años se entendió como una simple extensión de las artes dramáticas, según se concebían por el teatro de esa época. *El cinematógrafo era en un principio un simple medio de filmar obras teatrales* [O'Donnell, 1994]. Las expresiones, el ritmo y los enfoques se hacían más bien como obras de teatro simplemente filmadas a través del adelanto tecnológico representado por la cámara cinematográfica, la que, como un nuevo arte de la representación, como hemos podido constatar a lo largo de los años, tiene su propia personalidad y dinámica dentro de esas artes de la representación. Lo mismo le sucedió a la televisión: durante un buen tiempo era sólo una extensión de la radio. Los primeros años presentaron programas que seguían la dinámica de la radiodifusión, simplemente añadiendo una imagen, pero pasó un buen tiempo antes de que esa sociedad le imprimiera su propia personalidad.²⁰ Múltiples ejemplos pueden citarse al respecto: el automóvil en sus inicios, era construido

20 Voutssás M., Juan. 1995. "Crónica de Tres Muertes Anunciadas: El Teatro, la Radio y el Libro", en: *BIBLIOS: Revista para Especialistas en Información*, vol. I, no. 2, pp. 2-3. marzo-abril, 1995.

prácticamente igual que una carroza tirada por caballos, simplemente con un motor en su interior; los libros impresos durante las primeras décadas imitaban la escritura manuscrita, etcétera. Según O'Donnell, *toma de una a tres generaciones pasar del punto de concebir un nuevo medio a pensar realmente la manera de explotarlo **con gracia** en su propia esencia*. Debemos pues, tratar de ir más allá de esa inercia y tratar de plantear su nuevo contexto desde esta primera generación de las bibliotecas digitales, de concebir las con ésta, su propia personalidad, y de explotarlas *con gracia*.

Las bibliotecas en México, al igual que en otras partes del mundo no son iguales entre sí: existen bibliotecas públicas, universitarias, escolares, especializadas, etcétera. A sus acervos hay que sumar las diversas colecciones documentales, fotográficas, musicales, museográficas, etcétera, que existen en otras instituciones fuera de las bibliotecas; los archivos documentales históricos, los archivos cinematográficos y videograbados de muchas otras organizaciones. Éstas tienen diversos sistemas de registro, distintas comunidades usuarias y múltiples necesidades que deben cubrir. Agreguemos los nuevos elementos derivados del auge de la publicación electrónica y la *Internet* en los años recientes, los aspectos de formatos y estándares, los aspectos legales y tendremos un panorama de lo complejo que resulta tratar de definir un modelo de biblioteca digital para nuestro país.

De todo lo anterior, y tratando de hacer un resumen de todos los elementos para su construcción y bajo un enfoque reduccionista —por ahora—, he aquí una relación de los enfoques o aspectos que deben ser contemplados por los diseñadores en la conceptualización y puesta en marcha de un proyecto de biblioteca digital:

1.4 ENFOQUES

Para un correcto análisis de una biblioteca digital, es importante observarla desde varias perspectivas, o “enfoques”, los cuales permiten entenderla y diseñarla desde distintos puntos de vista complementarios uno con el otro:

Desde el enfoque de las colecciones:

- Definición del perfil de la colección; políticas de selección, preservación y descarte.
- Conversión a soportes digitales de documentos que se encuentran originalmente en medios tradicionales.
- Documentos originalmente digitales; publicaciones digitales.
- Establecimiento de grandes bibliotecas depositarias.
- Sustentabilidad de las colecciones a lo largo del tiempo.

Desde el enfoque de la tecnología:

- Arquitectura de la plataforma computacional.
- Criterios de digitalización; viabilidad de proyectos; recursos propios o contratados.
- Almacenamiento de las colecciones: costos actuales y proyectados. Almacenamiento en-línea y fuera de línea.
- Establecimiento de los formatos aceptados en la biblioteca para cada tipo de material digital y técnicas de compresión aceptadas.
- Recuperación de información. Diseño de las políticas computacionales de registro de los materiales para cada colección. Tipo de metadatos aceptados: Marc 21, MODS, Dublin Core, XML, Z39.50, etcétera. Planes de viabilidad del desarrollo de esos metadatos.
- Establecimiento de las Interfaces para usuarios aceptadas para los servicios digitales.
- Presentación de la biblioteca digital: el portal, página *Web* o sitio de la misma.
- Políticas y costos para la entrega de los contenidos y servicios.
- Estándares aceptados por la biblioteca, como por ejemplo: OEB, Z39, TCP/IP), etcétera.
- Políticas de interoperabilidad con otros sistemas.
- Modelos de continuidad del servicio; éste debe incluir planes de contingencia en caso de fallas.
- Manejo de versiones tecnológicas.

Desde el enfoque de los usuarios:

- Establecimiento de niveles de usuarios: hacia arriba —de quién somos usuarios—, mismo nivel —de quién somos pares—, hacia abajo —quiénes son nuestras nuevas comunidades de usuarios—.
- Características de los usuarios: profesores, investigadores, creadores, estudiantes, comunicadores, profesionales, distribuidores, gran público.
- Métodos de evaluación: necesidades, preferencias, comportamiento y habilidades de los usuarios, ya sean individuales, en grupo o gremiales.

Desde el enfoque de la organización documental:

- Recuperación de la información. Diseño de las políticas bibliográficas de registro de los materiales para cada colección
- Tipo de metadatos descriptivos aceptados: Marc 21, MODS, Dublin Core, Z39.50, etcétera. Planes de viabilidad del desarrollo de esos metadatos.
- Clasificación e indización, manual y automatizada.
- Metadatos administrativos, técnicos, descriptivos, de hipervínculo con entes exteriores al documento, etcétera.
- Lenguajes de marcado: HTML, SGML, XML, MODS, METS, etcétera.

Desde el enfoque legal:

- Establecimiento claro de los límites y prerrogativas que la ley autoral ofrece a esa biblioteca.
- Definición de las políticas de “uso ético de la información” —*fair use*— para esa biblioteca.
- Definición de material para *Intranets*.
- Costos de transmisión de derechos.
- Definición de políticas de copia para lo que la biblioteca produce por parte de la propia institución.

- Uso de definiciones externas para cesión de derechos de lo propio, *Creative Commons*.
- Establecimiento de criterios de privacidad por parte de la biblioteca.

Desde el enfoque social:

- Declaración de principios de la biblioteca respecto a la sociedad de la información o sociedad del conocimiento.
- Establecimiento de la relación acerca de nuevos modelos educativos, modalidades educativas y la biblioteca.
- Alfabetización digital; reducción de la brecha digital.
- Proyectos desprendidos de colaboración estrecha de la biblioteca digital con la educación.
- Derecho a la accesibilidad.
- Creación de contenidos por parte de la biblioteca digital.
- Modelo de sustentabilidad económica de la biblioteca digital; costos y financiamientos.
- Preservación de valores, identidad y cultura.
- El perfil deseado del personal profesional en esa biblioteca digital.
- Balance entre intereses de los productores y de los usuarios.
- Educación para el aprovechamiento de las bibliotecas digitales.

Algunos de estos puntos de vista se complementan, se superponen y se contraponen, y queremos tratarlos con cierta organización metodológica al momento de analizarlos, pero trataremos de incluir todos y cada uno de ellos durante esta etapa del estudio detallado de la biblioteca digital, previamente al planteamiento del modelo nacional. Cabe recordar que en esta primera parte queremos definir las partes que componen una biblioteca digital, *vista como un ente aislado*, con objeto de ir entendiendo sus partes e interrelaciones.

Una primera conclusión muy pertinente a este capítulo: Las bibliotecas digitales están contenidas dentro de la *Internet* y por tanto son parte de los recursos documentales de la red mundial. La *red* como un todo *no* es una Biblioteca Digital.

Capítulo 2

Colecciones digitales convertidas

“La Biblioteca existe ab æterno. De esa verdad cuyo colorario inmediato es la eternidad futura del mundo, ninguna mente razonable puede dudar”.

de la “La Biblioteca de Babel”, Jorge Luis Borges.

2.1 DESARROLLO DE COLECCIONES: CONCEPTO Y ANTECEDENTES.

Es obvio que una biblioteca digital ofrece a sus usuarios una colección digital, al margen de todo lo que pueda poseer en sus acervos tradicionales. Lo que ya no es tan obvio es cómo desarrollar una colección digital adecuadamente. Primeramente, es necesario establecer que, al igual que en una biblioteca tradicional, no existe una única colección documental. En realidad está formada por un número de colecciones parciales de distintos tipos de materiales y formatos que van conformando lo que la biblioteca y sus usuarios conocen como “la colección” en términos generales. Algo semejante sucede en una biblioteca digital; en realidad su oferta documental está formada por un conjunto de distintas colecciones, así como formatos y materiales diversos que se ofrecen al público en un formato digital. De esta forma, cuando me refiero a la “colección digital” de una biblioteca significa el conjunto de ellas que conforman esa colección.

Como toda colección bibliográfica, debe establecerse un objetivo y una política de desarrollo para la misma. Esta es una condición *sine qua non* y que sin embargo es soslayada frecuentemente, y se procede a crear una colección digital de forma directa. Al igual que sucede con una biblioteca tradicional, no puede desarrollarse adecuadamente una colección en el entorno digital sin haberse establecido previamente y claramente cuál es el objetivo que se persigue al crear esa colección, a quién va dirigida, cuál es su nivel, alcance, cobertura, política de adquisición, permanencia y descarte de materiales, etcétera. Aquí es donde la capacitación y la experiencia del bibliotecario profesional es sumamente útil para establecer esa política de manera adecuada y el

momento en que debe quedar por escrito en un documento para funcionar como guía de largo alcance en la biblioteca.

Una vez establecida la política de desarrollo de la colección es necesario establecer una tipología de los materiales que la conformarán. Para efectos de este trabajo utilizaremos esta forma de tipificar los materiales, es decir, según sus fuentes o forma de obtención. Hay varias formas de agrupar tales fuentes, pero en términos generales se reconocen tres grandes agrupamientos de donde los materiales para la colección digital pueden provenir:

- 1) Materiales en soportes tradicionales que la biblioteca posee y que son susceptibles de digitalizarse.
- 2) Materiales originalmente digitales, ya sea de la propia institución a la que la biblioteca pertenece o adquiridos a proveedores.
- 3) Hipervínculos a otras colecciones y materiales digitales que pueden ser asociados a nuestra colección, dando la percepción al usuario de que pertenecen a ésta.

El advenimiento de la tecnología digital como un nuevo soporte de preservación de material introdujo una serie de variables y circunstancias que deben ser consideradas para no crear un caos. En este sentido cabe mencionar que como primer antecedente en la Unión Americana existió el *Digital Preservation Consortium* —DPC—. Este organismo fue formado por un grupo de universidades de la unión americana que trabajan conjuntamente tanto a nivel de instituciones como de académicos para conseguir un uso efectivo de los materiales preservados en formato digital. La misión de este consorcio era desarrollar el uso y la utilidad de la tecnología digital para preservar e implementar el acceso a obras intelectuales de importancia nacional e internacional.²¹ Las universidades integradas a este consorcio fueron: La Universidad de California, la del Sur de California, la de Stanford y la de

21 Commission on Preservation and Access. (EUA). 1994. *The Digital Preservation Consortium Mission and Goals*. Disponible noviembre 3, 2005 en: <http://www.clir.org/pubs/reports/dpcmiss/dpcmiss.html>

Berkeley; las universidades de Columbia, Michigan, estatales de Pennsylvania y de Tennessee, así como Harvard, Yale, Cornell y Princeton.

Éstos son los objetivos que en 1994 se propusieron para los siguientes cinco a diez años, y que alcanzaron con mayor o menor éxito:

- Verificar y supervisar el uso de las imágenes digitales para su preservación y acceso.
- Definir y promover métodos y criterios compartidos, para la producción, almacenaje y distribución de imágenes digitales.
- Ampliar la base de materiales preservados en forma de imágenes digitales.
- Desarrollar y mantener mecanismos fiables que garanticen el acceso generalizado a documentos digitales de imágenes.

Este organismo se convirtió un poco más adelante en la *Digital Libraries Federation* o Federación de Bibliotecas Digitales, la cual inició la *Primera Iniciativa en Bibliotecas Digitales*, como ya se comentó anteriormente. De hecho, los resultados no fueron tan importantes como la misión presentada en ese entonces. No obstante, la DLF sentó las bases para proyectos de digitalización muy importantes en la unión americana. En la actualidad, los Estados Unidos han vuelto a desencadenar otro enorme proyecto cooperativo con miras a la preservación masiva y de largo término de su patrimonio documental electrónico, encabezado por la Biblioteca del Congreso de los EUA, a través de un programa denominado “*The National Digital Information Infrastructure and Preservation Program*” (NDIIPP). Disponible en: <http://www.digitalpreservation.gov/>.

2.2 CONVERSIÓN DE DOCUMENTOS ORIGINALMENTE EN SOPORTES TRADICIONALES

Cuando se piensa en bibliotecas digitales, si no se tiene en mente el consenso del Taller de Santa Fe, inmediatamente viene a la mente la idea de que están formadas, tautológicamente, por colecciones digitales. En la práctica puede observarse que están conformadas por un amplio

rango de materiales, tecnologías y prácticas híbridas, más allá de lo únicamente digital.

Las bibliotecas digitales del mundo están conformando sus colecciones hacia el mundo digital de dos maneras: adquiriendo acceso a materiales producidos ya en forma digital de origen, vía servicios de red o dispositivos de almacenamiento local, como CD-Rom, cartuchos de cinta, DVD, etcétera; o bien digitalizando colecciones propias que existían previamente en la biblioteca en algún soporte tradicional. De todos modos, es importante considerar que durante un buen tiempo coexistirán los materiales en soportes tradicionales con los materiales en soporte digital.

Antes de entrar en materia acerca de la construcción de acervos digitales, es conveniente hacer un breve análisis de algunas características propias del copiado de documentos, con objeto de poder entender mejor el contexto de las colecciones de ellos.

Los documentos pueden ser a la vez ser fijos y cambiantes; a través del uso de ciertas “tecnologías de fijado” somos capaces de inscribir ciertas marcas con significado en alguna superficie de modo tal que éstas puedan trascender el tiempo y el espacio, y permitirle así a muchas personas leer lo mismo al momento de fijarlas y por un cierto tiempo después. Pero afirmar que los documentos son “fijos” no significa que éstos no cambien y sean eternos; todos los documentos son “fijos” por un cierto periodo de tiempo y son cambiantes en otras épocas. De hecho, diferentes documentos tienen diferentes ritmos de fijado y cambio.

Hay documentos que aún existen en su soporte original hasta nuestros días. En otros casos, si el documento es lo suficientemente valioso, irá pasando por múltiples pasos de fijado y cambio a lo largo del tiempo. El documento original puede llegar a perderse, pero su contenido, sus “fijados” en algún soporte pueden trascender por generaciones y generaciones. Como ejemplo de los primeros tenemos a la famosa *Piedra de Rosetta*, erigida como una estela en el año 196 a.C., y los *Rollos de Khirbet Qumrán*, escritos en papiro cerca del siglo I a.C, los cuales pueden ser admirados hoy en día todavía en su soporte original. Como ejemplo de los segundos tenemos las tragedias de Esquilo, Sófocles y Eurípides. Nadie ha visto hoy en día una en su soporte original; ni

quiera en los recientes pasados siglos. Fueron escritas en el siglo V a.C., muy probablemente usando papiro como soporte, y tinta obtenida de sales de hierro, taninos de madera, hollín y goma, grafiando en *Ático*, uno de los idiomas griegos de esa época. Ése, o alguno muy parecido, debe haber sido el primer fijado. A lo largo de casi veinticinco siglos esas obras han sido cambiadas a muchas otras versiones de fijado; algunas de ellas cambiando su soporte; es decir, pasaron sucesivamente a pergamino, a papel de trapo y luego a papel de celulosa; fue cambiando también la herramienta para fijarlas; se usaron cálamos para copiarlas a mano y después imprentas tipo Gutenberg hasta llegar a las litográficas y a las computarizadas; cambió también el lenguaje en el que estaban escritas. Aún las versiones griegas actuales son sustancialmente distintas de las de ese entonces; el griego actual es tan parecido al griego antiguo como lo es al español. Deben haberse fijado y cambiado estos documentos docenas, cientos, y tal vez hasta miles de veces. Siguiendo una historia similar a las obras del ejemplo, muchas se perdieron en el viaje hasta nuestros días, pero por fortuna muchas otras han sobrevivido a su epopeya y son nuestras contemporáneas. Algunas de las obras de la antigüedad griega pudieron sobrevivir hasta nuestros días siendo copiadas al latín, de ahí al árabe, perdidas en Europa y redescubiertas después en esa lengua gracias a la árabiga *Escuela de Traductores de Toledo* en España o a la *Casa de la Sabiduría*, —*Bayt-al-Hikmah*—, la enorme biblioteca-centro de investigación del Califa Harun-al-Rashid en Bagdad quien copiara a todos los clásicos griegos y latinos del Imperio Bizantino de Oriente. Fueron recopiadas luego al latín y al español, y de ahí nuevamente a numerosas lenguas actuales. Hay varias razones por las que ha existido el deseo de realizar ese cambio y un nuevo “proceso de fijado”. Entre los principales factores están:

- Que el documento despierte suficiente interés para ser copiado al menos una vez.
- Que el documento despierte tal interés que se desee copiarlo más veces, incluso a otras lenguas.
- Que el nuevo soporte garantice un número razonable de años de existencia en esa nueva versión del documento.

- Que el documento logre un cierto número de copias que maximicen su probabilidad de sobrevivir o bien, que alguna copia llegue a algún lugar que se preocupe por preservarlo —por ejemplo una biblioteca—.
- Y finalmente, que el documento original tenga la suerte de sobrevivir una serie de destinos adversos, factor que como biblioteca-rios, obviamente, quisiéramos que siempre ocurriera.

Cuando se habla entonces de “convertir” los acervos de una biblioteca contenidos en soportes tradicionales, no estamos hablando por lo tanto de nada nuevo como concepto. Como puede verse, simplemente estamos realizando un proceso de fijado en un nuevo soporte, probablemente en una nueva lengua, y tal vez agregando o modificando algunas de sus características; esto es muy natural. Sin embargo, nos inclinamos a pensar que el cambio de un soporte tradicional actual, llámese papel, disco fonográfico de acetato o pasta, o película de celuloide hacia un soporte electrónico es tan radical que impone un nuevo orden de las cosas. Tendemos a olvidar que el papiro y la tinta china fueron utilizados por milenios antes de cederle el paso a sus sucesores; que el pergamino y la tinta actual fueron usados por bastante más de mil años; que el papel de celulosa del tipo que conocemos no llega a doscientos años y que la imprenta apenas rebasa los cinco siglos. Que los discos fonográficos y la película de celuloide o similares apenas frisan los cien años. Por supuesto, cada soporte tiene sus desventajas; no existe ninguno ideal: el papel se humedece, se pudre y se quema fácilmente; los microfilmes y soportes electrónicos obsolescen y requieren de dispositivos tecnológicos para leerse, etcétera.

Dejando a un lado entonces el elemento novedoso del cambio de soporte, puesto que no lo es, lo que realmente debe considerarse a la hora de pensar en convertir documentos a formatos digitales es el porqué se desea hacerlo. Existen múltiples razones por las que puede desearse convertir un documento, pero en realidad pueden resumirse en dos: preservar o distribuir.

2.2.1 Preservación y conservación

*“La información digital dura para siempre,
o por cinco años, lo que ocurra primero”.*

Jeff Rothenberg; científico ‘senior’
en la Corporación Rand.

Para definir estos conceptos conviene usar las sencillas pero efectivas definiciones de Wesley Boomgaarden, Director de la oficina de restauración de las Bibliotecas de la Universidad Estatal de Ohio, EUA.²²

Preservación: Acciones tomadas para anticipar, prevenir, detener o retardar el deterioro de obras”.

Conservación: El mantenimiento otorgado a cada pieza en la colección para tenerla en condición de uso.”

Restauración: Acción de regresar una pieza deteriorada a su condición original o cercana a ella.”

Haciendo a un lado la tarea de restauración, puesto que no viene a colación en esta temática, nos centraremos en las de preservación y conservación. El deseo de preservar tiene que ver obviamente con el deseo de heredarle un documento a futuras generaciones. Dependiendo de la fragilidad del soporte en el que se encuentra, de su edad y de su estado de conservación, el documento puede estar en riesgo de destrucción inminente, y debemos preocuparnos entonces de fijarlo en otro soporte para poder evitar, si no la destrucción del documento-soporte como se conoce, cuando menos la de su contenido. Tal sería, por ejemplo, el caso de los *Rollos del Mar Muerto*. En otros casos, el factor de preservación puede ser el hecho de evitar su manipulación, a pesar de que su estado de conservación sea satisfactorio, para alargar la vida del soporte actual. Tómese por ejemplo el caso de un incunable. Por lo general, si está más o menos conservado, su edad no lo pone en riesgo

22 Boomgaarden, Wesley (ed.). 1993, *Staff Training and User Awareness in Preservation Management*. Preservation Planning Program Guide 2. Washington, D.C.: Association of Research Libraries.

inminente de destrucción, pero sabemos que cada vez que ese libro se abre y cierra y que sus páginas son pasadas, lo sometemos a un desgaste físico el cual, si ello fuese frecuente, en algunas pocas décadas, el esfuerzo mecánico, la exposición a la luz, etcétera, lo deteriorarían inexorable y gravemente lo cual, por supuesto, no es deseable.

Más allá de la preservación se tiene, por parte de los bibliotecarios, el deseo de distribuir ampliamente la información; de que se divulgue, se desplace, se consulte; y de que esté al alcance del mayor número posible de lectores. Por ello, puede desearse entonces pasar esa información a un nuevo soporte, para maximizar las posibilidades de que ello suceda. Se busca un soporte ágil, que esté al alcance de muchos usuarios y, por supuesto, que esto pueda hacerse a costos *razonables*.

En estos casos no existe tanto la necesidad de preservar; al menos no tan urgentemente; hay soportes físicos lo suficientemente robustos como para no estar en riesgo inminente o de manipulación, como los documentos grabados en piedra. O puede tratarse de ediciones en papel lo suficientemente nuevas y numerosas para no tener que preocuparse todavía por la manipulación, pero el deseo de optimizar su distribución persiste.

Es ésta la primera decisión de fondo que el bibliotecario debe tomar al momento de pensar en digitalizar una colección, ya que de esa decisión dependen muchas cosas. En primer lugar conviene establecer que el resultado de la digitalización a nivel documental, debe ser lo que se conoce como *reproducción digital fiel*. Este concepto significa que el documento resultante reúne tres características indispensables: *calidad*, en el entendido de ésta como su funcionalidad y valor de uso; *permanencia*, en el concepto de que el documento podrá ser accesible por un lapso considerable; e *interoperabilidad*, lo que significa que el documento será accesible a lo largo de varias plataformas y programas de computadora. En lo tocante al primer elemento, la calidad, conviene abundar que esto puede incluir características tales como la completud del documento fuente, la apariencia del original —incluyendo tono y color—, su originalidad, escala uno a uno, etcétera.

Cuando a esta *reproducción original fiel* se le agregan los metadatos, es decir, los elementos que permiten su recuperación posterior, obtenemos un *objeto digital* u *objeto de información*. Objeto digital

es, por tanto, cualquier entidad documental: texto, imagen, sonido, etcétera, que ha sido codificada numéricamente bajo algún formato y ensamblada junto con algún conjunto de metadatos de tal forma que puede ser almacenada, buscada, encontrada, y usada a partir de una colección dentro de un sistema de cómputo.

Por tanto, la primera decisión a tomar, como ya hemos mencionado, es establecer si vamos a crear objetos digitales para preservar, para distribuir, o para ambas cosas, ya que como veremos más adelante las características y los costos de cada una de ellas son *sumamente distintos*. Nunca debemos emprender proyectos de preservación con especificaciones de distribución ni viceversa. Las consecuencias pueden ser lamentables.

Una vez que se han establecido los dos fines primordiales por los que una biblioteca podría querer digitalizar una colección de materiales ya existentes en soportes tradicionales, cabe destacar entonces que de acuerdo al consenso de los principales grupos de especialistas a nivel mundial, después de varios años de observación del desarrollo de esta tecnología, sus capacidades, costos, etcétera, el consenso es que *la digitalización es una excelente forma de maximizar el acceso a materiales; es decir, con fines de distribución es una formidable herramienta* y debe ser usada ampliamente. Pero si se trata de preservar documentos a largo plazo la historia es otra; el contexto tecnológico todavía no está del todo maduro; puede usarse *con cautela*, y procurando tener algún respaldo de otro tipo; por ejemplo, se recomienda digitalizar a partir de copias microfilmadas de documentos y no viceversa. Y lo mismo cuando el estado físico de los documentos originales sea sumamente grave. En las palabras de Abby Smith:²³ *“digitalización no es preservación; mucho se gana digitalizando, pero permanencia y autenticidad, en el actual estado del desarrollo tecnológico, no están entre esas ganancias”*. Una opinión parecida puede

23 Smith, Abby. 1999. *Why Digitize?* CLIR. Council on Library and Information Resources. Commission on Preservation and Access. 2:99, ISBN 1-887334-65-3. Disponible noviembre 3, 2005 en: <http://www.clir.org/pubs/reports/pub80-smith/pub80.html>

verse en el Reporte Final del Grupo de Trabajo de la Asociación Alemana de Investigación.²⁴

2.3 CRITERIOS PARA DIGITALIZAR

Como hemos podido observar, los acervos digitales facilitan en mayor o menor grado el cumplimiento de estos dos propósitos, preservar o distribuir, pero es necesario establecer previa y claramente criterios para hacer esto. Frecuentemente se piensa en digitalizar todo. Si bien la idea es atractiva de inicio, presenta formidables obstáculos para ser llevada a cabo. El número de documentos producidos desde el advenimiento de la imprenta acumula ya muchos millones de ellos. Aún sin pensar en el todo, una biblioteca con un acervo importante tendría enfrente la tarea de convertir millones y millones de páginas, lo cual impone enormes recursos y obstáculos de todo tipo que no hacen fácil la empresa.

Es una idea comúnmente aceptada, aunque no siempre bastante reflexionada, la de que construir bibliotecas digitales consiste simplemente en digitalizar los documentos que ya se tienen en otros soportes. Ello conlleva otros riesgos; la inversión requerida para digitalizar textos es muy considerable. Por ejemplo, hablando de textos, el costo total se acerca en la actualidad a un dólar por imagen convertida. En colecciones de magnitud significativa la inversión puede resultar demasiado desgastante para bibliotecas que de entrada tienen remanentes muy marginales para invertir en este tipo de proyectos. La conversión digital en un ambiente operacional requiere de voluntades políticas institucionales de largo plazo, una integración tecnológica considerable y personal muy avezado en los estándares y manejo de este tipo de información.

Partiendo entonces del hecho de que no es recomendable digitalizar toda nuestra colección, sino sólo algunas partes de ella, es importante

24 Weber, Hartmut, y Dörr, Marianne. 1997. *Digitization as a Method of Preservation?*, CLIR. Council on Library and Information Resources. Commission on Preservation and Access. 10:97. ISBN 1-887334-53-6. Disponible noviembre 3, 2005 en: <http://www.clir.org/pubs/reports/digpres/digpres.html>

establecer *perfectamente* cuáles deben ser los criterios para que una organización, entre ellas la biblioteca, emprenda un proyecto a este respecto. La pregunta entonces es ¿qué debe digitalizarse, y con qué prioridades? ¿cómo debe hacerse? ¿por qué queremos hacerlo? Para contestar estas preguntas conviene revisar lo que ya se ha establecido acerca de ellas en algunas de las bibliotecas que ya han emprendido este tipo de tareas y que se consideran *buenas prácticas*. Los criterios aprobados por una biblioteca deben aplicarse siempre a todas y cada una de las partes de la colección susceptibles de ser digitalizadas.

Puede comenzarse con alguno de los más simples; por ejemplo, los *Criterios de selección para la preservación por reformateado digital* emitidos por la *División para la preservación por reformateado* de la Biblioteca del Congreso de los EUA. [LC Preservation, 1999]. En ellos se menciona:

La selección de materiales para la preservación por reformateado digital se basa en el valor, uso, condición, características del documento original y conveniencia de las reproducciones digitales para uso y acceso.

- Valor. Se dará prioridad a los materiales de interés nacional, de alto valor y en riesgo. La entrega de copias digitales de este tipo de información es altamente deseable; en parte para servir como preservación preventiva, así como por su seguridad, lo que se pretende lograr al reducir el manejo de los originales.
- Condición. Los documentos que no estén en servicio por daño o fragilidad son candidatos para el reformateado digital, ya que son documentos que se encuentran actualmente en soportes inestables.
- Uso. Los materiales originales que tengan una tasa de demanda muy alta o altos costos de recuperación son fuertes candidatos para reformateado digital.
- Características del original. Los originales en distintos formatos físicos y con características diferentes son elegibles para reformateado digital, incluyendo documentos encuadrados o sueltos, fotografías, negativos, y documentos a color. Esta lista de posibles candidatos se irá expandiendo en un futuro.
- Aceptabilidad del objeto digital resultante. El objeto digital resultante del proceso de reformateado digital debe cumplir los requerimientos de la división de custodia y la Dirección de Preservación de la Biblioteca del Congreso. En algunos la tecnología y/o metodología para realizarlo puede no existir, o estar inmadura para una

producción masiva. En estos casos, algunos elementos de la colección o colecciones pueden ser elegibles como prueba para desarrollo de esos elementos de tecnología y/o metodología para futuros trabajos de reformateado.

- Ayudas al acceso. Las ayudas para el acceso de todo tipo son elegibles para el reformateado digital, especialmente cuando las ventajas de la tecnología digital signifiquen mejora para el acceso a colecciones en papel o microfilmadas. Por ello, las ayudas al acceso tales como índices, resúmenes, bases de datos, etcétera, son consideradas adecuadas bajo este propósito.

A continuación, analicemos los de simplicidad media; por ejemplo, los criterios de selección para digitalizar emitidos por la Universidad de California. En ellos, [Thompson, 1998] declara de inicio:

La siguiente lista de criterios es recomendada tanto para aquellos bibliotecarios que se dedican al desarrollo de colecciones como a los encargados de la preservación al momento de seleccionar materiales análogos para su conversión a formato digital. Algunos de los criterios están basados en la selección convencional y consideraciones de preservación comunes a todo formato; otras surgen de las oportunidades y características propias de las tecnologías digitales.

La colección a ser convertida y la tecnología digital a ser utilizada deben ser evaluadas contra algunos o todos de los criterios enunciados, dependiendo del grado de énfasis en el desarrollo de colecciones, acceso y objetivos de preservación perseguidos. Nótese que las tecnologías digitales actuales no se han estabilizado adecuadamente para la preservación de materiales cuando la versión digital se entiende como el remplazo, más que el suplemento de esa versión original.

A. Criterios aplicables a la colección que será convertida.

Desarrollo de colecciones:

1. Cumple las necesidades de información de estudiantes, profesores e investigadores universitarios dentro y fuera de la Universidad de California
2. Tiene opinión favorable para la conversión por alguna o algunas comunidades académicas significativas

Preservación / Archivo:

3. Tiene algún problema de preservación actual, tales como riesgo de daño por su uso, pobre almacenamiento, o deterioro físico visible.

Organización / Financiamiento:

4. Cumple los intereses del organismo patrocinador
5. Produce prestigio a la institución

B. Criterios aplicables a la versión digital.

Desarrollo de colecciones:

6. Agrega valor a la información por encima de las copias sobre papel, tales como mayor valor intrínseco a la información, mayor funcionalidad gracias a hipervínculos o mayor facilidad de búsqueda y navegación.
7. Contribuye a la 'Masa Crítica' de materiales digitales acerca de un tema y por lo tanto incrementa el valor de la colección.

Preservación / Archivo:

8. Crea aceptación por parte del bibliotecario para continuar manteniendo la colección a través del tiempo, cambios tecnológicos y riesgos posibles en la seguridad inherentes a archivos digitales.
9. Puede digitalizarse adecuadamente sin causar daño a originales, y le permite al documento digital fungir como sucedáneo del original, reduciendo así su uso y por tanto deterioro.

Acceso:

10. Puede ser integrado a los programas regulares de servicio bibliotecario.
11. Es accesible desde las plataformas computacionales y de red de la institución y cumple con los estándares aceptados por la comunidad bibliotecaria.
12. Puede ser entregado a velocidades razonables.

13. Resuelve problemas técnicos de acceso a los originales, como fragilidad, formatos voluminosos y/o incómodos, e integra formatos híbridos.
14. Se pueden instrumentar las restricciones al acceso necesarias por el tipo del material.
15. Puede utilizarse con interfaces entendibles por el usuario promedio.
16. Tiene o pueden agregársele metadatos adecuados para identificación, captura, origen, navegación, etcétera.
17. Puede ser autenticado.

Organización / Financiamiento:

18. Cumple los intereses del organismo patrocinador.
19. Es congruente con otras iniciativas de digitalización nacionales o gremiales.
20. Tiene comerciabilidad en su forma digital y/o crea una sociedad con algún proveedor de información”.

Finalmente, como ejemplo de criterios muy completos la Universidad de Columbia dice lo siguiente en lo tocante al establecimiento de criterios para digitalización [Columbia University, 2001]:

La creación de versiones digitales de impresos, fotografías, materiales de archivo, etcétera, puede ser llevada a cabo debido a varios propósitos, desde la investigación hasta el entretenimiento popular, y desde proyectos a gran escala hasta la digitalización de algunas imágenes para uso individual. Es caro seleccionar, crear y mantener recursos digitales; el costo de la captura de imágenes representa sólo un tercio del gasto total. Es por tanto importante asegurarse durante el proceso de selección que factores como la factibilidad técnica, los derechos de propiedad, y el apoyo institucional están salvados y son congruentes con el valor de los materiales y el interés sobre los contenidos. Los criterios enunciados no son exhaustivos ni obligatorios, y su importancia relativa variará con el propósito de cada proyecto. Están diseñados como ayuda a los planificadores para meditar cuidadosamente sobre los aspectos relevantes de este proceso.

El valor del contenido de los materiales y los beneficios derivados del acceso a las versiones digitales deben justificar la inversión de tiempo y es-

fuerzo del proyecto de digitalización. Los contenidos deben tener suficiente valor intrínseco para asegurar el uso a lo largo de un tiempo razonable por parte de la comunidad dentro y fuera de la Universidad de Columbia.

Valor (del contenido, físico, histórico):

Rareza o unicidad.

Valor como artefacto o asociativo.

Importancia para el entendimiento del tema del área.

Amplitud de cobertura de un área.

Contenido útil y preciso.

Información sobre temas pobremente documentados en otros medios.

Refuerza las fortalezas históricas de la institución.

Potencialmente duradero en forma digital.

Demanda:

Apoyo para actividades actualmente prioritarias.

Importancia para el funcionamiento de la institución.

Un público actual y activo para ese material

Aceptación del proyecto por la comunidad de Columbia.

Expectativa realista de atraer a nuevos usuarios.

Petición de socios potenciales para esfuerzos colaborativos.

No-duplicación:

No hay productos digitales semejantes que cubran las necesidades.

Como criterios de valor agregado podemos distinguir:

Potencial de colaboración:

Parte de la colección física está dividida entre varias instituciones que al ser digitalizadas pueden unirse.

Contribuye al desarrollo de una 'Masa Crítica' de materiales digitales en el tema.

Contribuye a la investigación en-línea.

Integra conjuntos de formatos o de materiales relacionados desperdigados.

Optimización del acceso académico:

Optimiza el acceso académico creando nuevas ayudas, ligas o desarrollo de índices.

Puede buscarse con facilidad, entre varios formatos y materiales.

Ayuda a diseminar ampliamente colecciones locales.

Optimización de la calidad de la imagen:

Mejora la calidad de imágenes previamente dañadas.

Criterios de propiedad intelectual:

El material es del dominio público, o la Universidad es el titular de los derechos.

El propósito de la colección digital cae dentro del ‘uso ético’.

Se puede llegar a un acuerdo con el titular de los derechos.

El acceso al material digital puede ser controlado o restringido.

Criterios de preservación:

Se reduce el manejo de materiales frágiles.

Se tiene acceso a materiales que es difícil usar de otra forma.

Se protegen materiales con riesgo de robo o mutilación.

Criterios de seguridad de creación:

La condición de los originales permite digitalizarlos con seguridad.

La condición de los originales requiere tratamiento y/o financiamiento para su digitalización.

Es posible digitalizar fotografías o equivalentes en lugar de los originales.

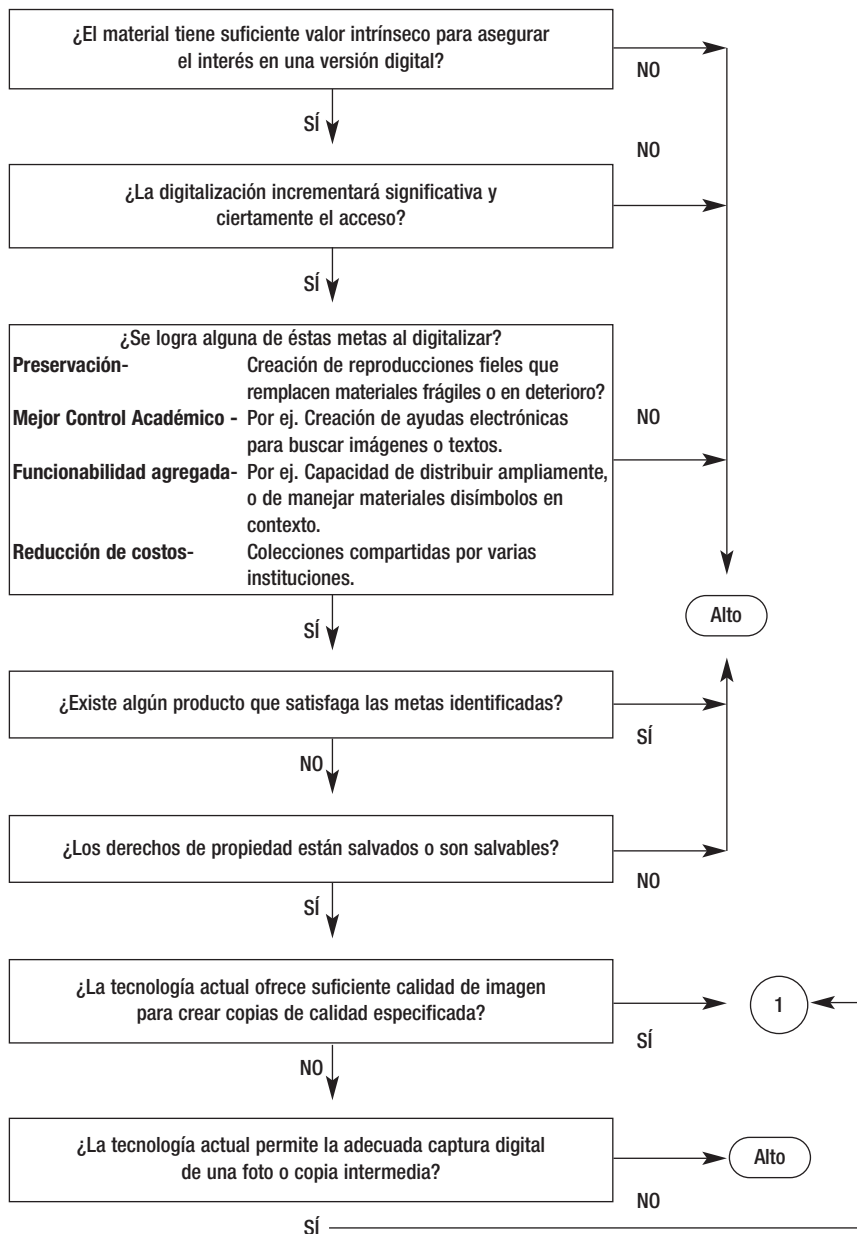
Criterios de viabilidad técnica; estos criterios incluyen:

Grado en que la versión digital puede representar el contenido total del original.

Comprensión de cuánta gente usará la versión digital y qué nivel de calidad de imagen implica.

Despliegue correcto de los materiales digitales.
Capacidad de acceder a la versión digital por parte de la plataforma usada en promedio por la comunidad.
Anticiparse a futuros y mejores equipos y por ende calidades.
Los materiales con especiales consideraciones incluyen:
Materiales que requieren resolución inusualmente alta.
Materiales en los cuales la fidelidad cromática es esencial.
Materiales de grandes dimensiones.
Objetos tridimensionales.
Materiales con poca legibilidad.

Como puede observarse, esta lista de criterios de Columbia es una de las más completas y que abarca más aspectos a considerar al momento de pensar en la digitalización de una colección. De acuerdo con el consenso general, su mayor aportación son sus criterios sobre “valor agregado”. La Universidad de Harvard, presenta estos criterios en forma de un diagrama de flujo, al cual denominan *Matriz de toma de decisiones para Digitalizar*: [Harvard University, 2001].



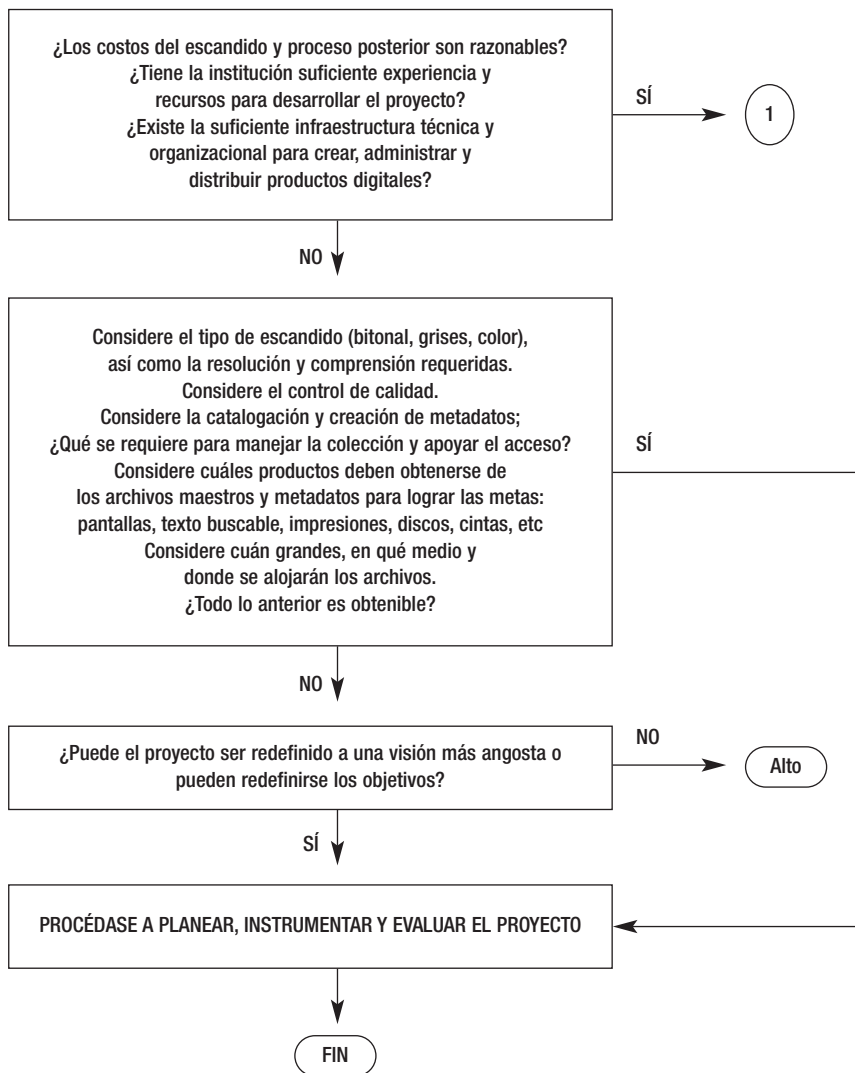


FIGURA 2.1 - “MATRIZ DE TOMA DE DECISIONES PARA DIGITALIZAR”. UNIVERSIDAD DE HARVARD.

Finalmente, y a guisa de resumen, presento la lista de los veintidós principios que deben regir el desarrollo de una colección digital. Éstos a su vez se subdividen en cuatro grupos: colecciones, objetos digitales, metadatos y proyectos de colecciones digitales [Baca, 2000]

Principio de las colecciones

- 1: una buena colección digital se crea de acuerdo con una política explícita de desarrollo de colecciones que ha sido acordada y documentada antes de que la colección comience.
- 2: la colección debe ser descrita de tal forma que el usuario pueda descubrir sus características, tales como el alcance, formato, restricciones al acceso, propiedad, y cualquier otra información significativa que determine la autenticidad, integridad e interpretación de la colección.
- 3: una colección debe ser sustentable a lo largo del tiempo. En particular, aquellas construidas con financiamientos especiales, las que deberán tener un plan que permita su utilización permanente más allá del periodo que abarquen los fondos.
- 4: una buena colección debe estar ampliamente disponible y evitar impedimentos innecesarios para su uso. Esto incluye accesibilidad adicional para personas con discapacidades así como su uso a través de tecnologías variadas y adaptables.
- 5: una buena colección respeta los derechos de propiedad intelectual. Los administradores de la colección deben llevar un preciso registro de los propietarios de los derechos y de los permisos aplicables a todos los materiales.
- 6: una buena colección tiene mecanismos para proveer datos acerca de su uso, así como otros datos que permitan mediciones estandarizadas acerca de su utilidad y demanda.
- 7: una buena colección encaja dentro de un contexto mayor de bibliotecas digitales nacionales e internacionales.

Principio de los objetos digitales

- 1: un buen objeto digital debe ser producido de tal manera que apoye las prioridades de la colección al tiempo que mantiene las cualidades que contribuyen a su interoperabilidad y reusabilidad.
- 2: un buen objeto digital es permanente. Esto es, debe existir la intención de una persona o institución identificados, de que ese objeto permanecerá accesible a lo largo del tiempo, al margen de tecnologías que cambien.
- 3: un buen objeto está representado en un formato digital que soporta el uso actual para el que fue diseñado así como usos futuros, o sea susceptible de ser copiado a otros formatos que sí lo soporten. Por tanto, un buen objeto digital es intercambiable entre plataformas, se hizo bajo algún estándar o buena práctica reconocidos y sólo se desvía de ellos debido a causas bien explicadas y documentadas.
- 4: un buen objeto digital será nombrado con un identificador único y permanente, de acuerdo con un esquema bien documentado. No deberá ser nombrado con base en una referencia absoluta a su nombre de archivo o dirección —como el caso de su URL— ya que éstos tienden a cambiar con el tiempo. En lugar de ello, un identificador estable debe poder ser resuelto o mapeado siempre a su dirección actual.
- 5: un buen objeto digital debe poder ser autenticable en al menos tres aspectos. Primero, el usuario debe poder ser capaz de determinar el origen, estructura, versión, y porción del objeto. Segundo, el usuario debe poder determinar que el objeto es lo que pretende ser. Tercero, el usuario debe poder determinar que el objeto no está corrompido o alterado de forma intencional.

Principio de los metadatos

- 1: los buenos metadatos son apropiados para los materiales de la colección, para sus usuarios, y deben pretender el acceso actual y futuro del objeto digital.

- 2: los buenos metadatos funcionan dentro de la interoperabilidad.
- 3: los buenos metadatos se basan en catálogos de autoridad y estándares de contenidos tales como vocabularios controlados, están alineados con las necesidades de los usuarios y describen adecuadamente tanto el contenido de los objetos digitales como sus objetos relacionados.
- 4: los buenos metadatos incluyen una expresión clara de las condiciones y términos de uso del objeto digital, incluyendo los legales.
- 5: los buenos metadatos perduran a largo plazo dentro de las colecciones.
- 6: los buenos registros de metadatos se convierten a su vez en buenos objetos digitales —tales como las bibliografías—. Por ello, deben tener las cualidades de los buenos objetos digitales, incluyendo autoría, autenticidad, archivabilidad, permanencia, e identificador único.

Principio de los proyectos

- 1: un buen proyecto de desarrollo de una colección tiene considerable diseño y planeación.
- 2: un buen proyecto de desarrollo de una colección digital incluye un plan de evaluación.
- 3: un buen proyecto de desarrollo de una colección digital produce reportes del proyecto y disemina adecuadamente la información derivada de cada etapa y los resultados del proyecto.
- 4: un buen proyecto de desarrollo de una colección digital considera el ciclo total de vida de la colección y sus servicios asociados a todo lo largo del proyecto.

2.4 FORMATOS DIGITALES

Hasta aquí hemos analizado algunas consideraciones básicas acerca del porqué de la digitalización, así como algunos de los criterios para iniciar este proceso. Es un buen momento para definir el término, ya

que puede crear confusión si no queda perfectamente establecido ¿qué significa exactamente digitalizar algo?

Para evitar caer en definiciones tautológicas, definiremos para comenzar el término *digital*. Éste fue utilizado por primera vez en 1938 al hacer una descripción de un circuito de cálculo que operaba datos en forma discreta, es decir, no continua: *El emisor... difiere de otros emisores en que tiene doce puntos conductores digitales*.²⁵

Según la definición del FOLDOC —“*Free On-Line Dictionary of Computing*”—²⁶ observamos:

Digital: dicese de una descripción de datos la cual es almacenada o transmitida en una secuencia de símbolos discretos provenientes de un conjunto finito; normalmente para este conjunto se usa información ‘binaria’ representada usando señales electrónicas, electromagnéticas, ópticas, etcétera.

Clarificando entonces, *digital* significa que una entidad de la naturaleza, es decir, un “objeto” está representado por “dígitos”; esto es, por “números”. Si bien en teoría podría utilizarse cualquier sistema numérico, en la práctica se utiliza casi siempre el sistema numérico binario, ya que es el ideal para ser utilizado en equipos electrónicos u ópticos. El sistema numérico binario se basa en sólo dos números, el “cero” y el “uno”, 0 y 1, los cuales pueden ser representados en forma electrónica —*existe corriente eléctrica = 1; no existe = 0*—; electromagnética —*existe magnetismo = 1; no existe = 0*—; óptica —*existe luz = 1; no existe = 0*— entre muchas otras formas de señalización física, lo cual hace más simple la construcción de ese dispositivo.

De esta forma, un documento digital es aquél que ha sido creado de origen o convertido a una forma de representación basada en números bajo un cierto patrón arbitrario, con objeto de poder ser almacenado o transmitido por medio de dispositivos electrónicos. *Digitalizar*

25 Citado en el *Timeline of Digital Preservation*. Entrada por el año 1938. University of Cornell. Disponible noviembre 3, 2005 en: <http://www.library.cornell.edu/iris/tutorial/dpm/timeline/index.html>

26 FOLDOC. *Free On-Line Dictionary of Computing*. Disponible noviembre 3, 2005 en: <http://foldoc.doc.ic.ac.uk/foldoc/index.html>

por tanto, es el proceso de convertir a un documento en *digital*, con objeto de poderlo almacenar y/o transmitir en una computadora, equipo de teleproceso, etcétera. Es decir, es un nuevo proceso de fijado como ya establecimos antes, sólo que en este caso este proceso significa convertirlo a un formato basado en números, generalmente binarios, bajo un cierto patrón arbitrario preestablecido antes de ponerlo en su nuevo soporte, un dispositivo asociado a una computadora.

Cabe resaltar que en esta definición, el concepto de “documento” debe considerarse en su acepción más amplia. Documento es entonces un texto en hojas sueltas o encuadernado como libro, sea manuscrito, mecanografiado, impreso, fotocopiado, etcétera. Es también una imagen fija, sea fotografía, dibujo, grabado, mapa, plano, microfilme, etcétera. Es también una grabación de sonido: disco, cinta magnetofónica, casete, etcétera. Documento es también una imagen en movimiento: película o video, etcétera.

Dado que *digitalizar* un documento es un proceso de representación del mismo por un conjunto de números, hay más de una representación posible para cada documento; es decir, pueden establecerse diversas maneras o “patrones” para digitalizar un documento. A un tipo de representación dada bajo una cierta convención arbitraria preestablecida se le conoce como un “formato”. Hay tantos formatos posibles como convenciones numéricas puedan establecerse para un documento —esto es, infinitas—. Muchas de estas convenciones tienen que ver con ciertas características deseables para ese tipo de documentos. No todos los formatos tienen las características ideales para representar adecuadamente a un documento. Depende de las necesidades y propósitos de quien lo crea.

2.4.1 Texto

Entre los documentos digitales de tipo “texto”, se distinguen dos tipos principales: los que son textos “buscables” y los que son textos “pictográficos”.

Los textos buscables son aquellos que están basados en el código ASCII, o en alguno de sus descendientes, como el UNICODE. En estos

códigos, cada letra o símbolo de escritura dentro del texto tiene un valor numérico único, estandarizado e internacional. Por ejemplo, cada una de las letras E, É, e, é, è, ê, ë, tiene un valor numérico único, distinto y universal. Cuando “tecleamos” un texto en una computadora y lo guardamos bajo un formato de archivo de texto, como los derivados de un procesador de palabra, estamos utilizando una variante del original código “ASCII”. Un archivo entonces con formato “TXT” —texto estándar— formato “DOC” —de “Word”—, formato “WPD” —de “word perfect”—, o formato “SWX” —de “Staroffice”—, son archivos de texto creados bajo un formato variante arbitrario al efecto. Su característica principal es que estos textos, al estar almacenados en una computadora, y al ser *todos y cada uno de sus caracteres diferenciables el uno del otro*, gracias a su valor numérico distinto, son susceptibles de ser analizados por algún programa, a fin de saber si un texto dado se encuentra contenido en alguna parte de cierto documento. Las bases de datos se crean también con variantes de este tipo de textos; gracias a esta característica es posible encontrar información contenida en ellas, y por ello podemos buscar y encontrar un autor, título, tema, etcétera, en estos archivos.

Los textos “pictográficos” son aquellos que han sido contruidos con un dispositivo que toma imágenes de la página en forma conjunta; por ejemplo, un escáner o una cámara digital. Cuando ponemos una página de texto en un escáner y la procesamos, el resultado es simplemente el escandido de la imagen en una serie de puntos consecutivos posicionados a lo largo de un plano. A este tipo de imagen electrónica se le conoce con el nombre genérico de *bitmap* —mapa de bits—. Si bien el archivo resultado es digital, y por tanto es también una representación numérica de esa página, sólo estamos posicionando puntos de colores a lo largo de un plano, pero para una computadora, son sólo puntos contiguos, y por tanto ella no puede saber realmente cuáles letras y palabras están ahí contenidas; es decir, es una imagen pictográfica. Para la computadora, *ese archivo está formado sólo por puntos, no por letras*. Nosotros los seres humanos podemos ver esas imágenes formadas por puntos y “reconstruirlas” dada la forma de cada letra, ya que esa es la manera en que generalmente las identificamos. Por lo tanto, para nosotros da lo mismo una página que fue tecleada y donde cada caracter tiene un valor numérico, o una página

de “imagen pictográfica” escandida a través de un escáner. Pero para una computadora ese texto es sólo una imagen sin significado alguno. Para ella realmente no es un texto, es sólo una serie de puntos. No puede buscar ni identificar nada dentro de él.

Esto no significa que ese texto no pueda utilizarse. Si lo deseamos, podemos agregar algunos elementos “tecleando” palabras clave en código ASCII que nos ayuden a la recuperación. De esta forma, si a un documento escandido “pictográfico” le agregamos una “ficha”, metadato o algún otro tipo de descriptor “buscable” tecleando estos datos, el texto “pictográfico” puede convertirse, aunque sea parcialmente, en texto “buscable” por esos elementos registrados; por ejemplo autor, título, tema, etcétera. Digo parcialmente porque el resto del texto del documento que no haya sido tecleado no puede ser buscado, pero muchas veces el hecho de contar con sus descriptores principales es suficiente. De hecho, se han recuperado libros en las bibliotecas por más de cien años sin tener el texto completo a la vista, gracias a una ficha catalográfica que contiene los datos precisos para ubicar esa obra.

Si se considera necesario, todo texto que haya sido escandido como imagen puede ser convertido a texto “buscable” mediante el proceso llamado “OCR” —*Optical Character Recognition*—, reconocimiento óptico de caracteres: proceso mediante el cual un programa de computadora creado para el efecto escande o “barre” un archivo de texto pictográfico tratando de interpretar cada uno de los puntos que forman caracteres, de la misma forma que lo harían nuestros ojos y cerebro con objeto de convertir ese texto desde la serie de puntos a texto formalmente hablando, asignándole a cada carácter su valor numérico correspondiente y convirtiendo esa página en una igual que si hubiese sido tecleada. De esta forma, en teoría, todo el archivo de imagen puede volverse texto buscable.

Este proceso sólo debe intentarse cuando sea absolutamente necesario. Su efectividad no ha sido optimizada del todo a lo largo de los años durante los que esta técnica ha existido. En la práctica, su éxito depende de una serie de factores que facilitan u obstaculizan dicha tarea. ¿Cuáles son en esencia esos factores?

- Nitidez y contraste del original —¿son originales o copias?, ¿es texto escandido, impreso o mecanografiado?, ¿sigue nítido?, ¿está manchado?, etcétera—.
- Defectos de origen —ejemplo: curvatura del texto cercano a la encuadernación—.
- Tipos de fuentes de letras distintas en los textos que van a ser interpretados.
- Tamaños distintos de las letras que contienen los textos.
- Cantidad de diacríticos que contiene el texto —¿cuántos idiomas y por lo tanto diferencia en los alfabetos hay involucrados en el texto?—.
- Caracteres no reconocidos —símbolos matemáticos, subíndices, etcétera—.
- Abundancia de secciones que no son propiamente texto: tablas, gráficos, diagramas, etcétera.
- Época de las fuentes de las letras —tipografía muy antigua, descontinuada hace muchos años, o tipografía gótica alemana—.

Si estos factores se combinan en forma adversa pueden hacer el proceso de interpretación tan lento y pobre, que puede llegarse a la conclusión de que es mejor y más rápido volver a teclear todo el texto que pretender interpretarlo con OCR. En teoría, un programa de este tipo puede llegar a interpretar hasta un manuscrito, ya que por lo general van “aprendiendo” de los textos para mejorar su nivel y calidad de interpretación, pero en la práctica pueden resultar muy ineficientes. En otros casos, después de algunos intentos, el proceso se vuelve muy estable y pueden interpretarse enormes cantidades de páginas rápida y certeramente. Lo mejor es siempre hacer una prueba previa, asegurándose de que la muestra a digitalizar *es representativa de toda* la colección a interpretarse y no tan sólo de una parte mínima. Los resultados de la prueba decidirán si se intenta el todo o se cancela la interpretación. Una versión mucho más profunda de las características de trabajo con OCR es presentada por Haigh, [1996].

Los textos digitalizados, por lo tanto, coexisten en una biblioteca digital en sus dos vertientes: buscables y pictográficos. Por lo general se observa que en ellas existe una buena cantidad de páginas “pictográficas”

digitalizadas por medio de un escáner o cámara, y que una cierta cantidad de esa información ha sido “tecleada” o convertida como texto buscable, para ayudar a su acceso. Sin que sea una división absoluta, por regla general en la información digital proveniente de acervos “convertidos” predomina la de tipo pictográfico, con una proporción menor de texto buscable, mientras que en la información proveniente de textos originalmente creados en forma electrónica sucede lo contrario.

¿Cuáles son los formatos más utilizados en el manejo de textos dentro de las colecciones digitales? Si bien no existe ningún estándar de norma a propósito, se observan algunos estándares *de facto* o recomendaciones. Por regla general, como ya se ha mencionado, existen combinaciones de dos o más formatos, dependiendo de qué tipo de texto se tiene representado. Así, para la publicación de textos en la red mundial, la *www*, se observa que el formato más utilizado es el HTML, —*Hyper Text Markup Language*— al cual me referiré con más detalle posteriormente. En cuanto a los formatos utilizados para la publicación electrónica más formal, por parte de los editores al respecto, se observa que el más utilizado es el formato “PDF”, —*Portable Document Format*— proveniente de la compañía “Adobe”. Después del HTML, el formato PDF es sin duda el formato más difundido en *Internet*. Este formato, surgido en 1993, es una variante del *Postscript*, lenguaje de programación diseñado para la definición de páginas para sistemas de impresión de alta definición. Por esta razón, los archivos PDF pueden incluir gráficos vectoriales, de tal forma que al ampliar la imagen no se pierda calidad. La característica más importante que presenta es que conserva el aspecto original de un documento. Además, permite realizar índices, hipervínculos y buscar palabras dentro del mismo documento. Es un formato idóneo para la impresión de documentos. Prácticamente todos los formatos se pueden convertir directamente a formato PDF. El programa sabe distinguir entre lo que es imagen y lo que es texto. Además conserva los hiperenlaces del documento. Un caso especial y que sólo se utiliza cuando se quiere digitalizar colecciones retrospectivas, es el formato *PDF imagen*. Como su nombre lo indica, se trata de un formato imagen y por tanto, no es posible buscar términos, ni conservar hipervínculos. Este caso es utilizado cuando se digitaliza una colección

retrospectiva de una revista o libros antiguos. Una vez obtenida la imagen a través del escáner, el documento se puede archivar como PDF.

Las principales ventajas de este formato consisten en que:

- Conserva el aspecto original del documento
- Puede contener múltiples páginas en un solo archivo
- Existe la posibilidad de usar el *zoom*
- Incluye diversas opciones de desplazamiento: miniaturas, marcadores
- Posibilita la creación de hipervínculos
- Permite la vinculación con multimedia
- Permite la creación de Formularios
- Incluye opciones de seguridad
- Tiene texto reconocible y editable (incluye módulo de OCR)
- Incluye la posibilidad de incorporar comentarios y notas

El PDF permite representar textos de una manera muy dinámica, con variedad de tipos de letra, distribución física en la página, inserción de gráficos, etcétera. Para los ambientes internos que se dedican a la producción de textos profesionalmente, tales como editores de libros, revistas o periódicos; agencias de publicidad, etcétera, los más difundidos sin duda son “QuarkXpress” de la empresa “Apple” y “Pagemaker” también de “Adobe”.

Para la representación “buscable” de esos textos, se agregan con frecuencia documentos de tipo “TXT”, llamado también “texto plano” el cual, siendo muy simple y limitado en sus posibilidades de despliegue, es enormemente transportable. Se ven también archivos provenientes de procesadores de palabra, tales como los de formato “DOC”. Mucha información “tecleada” es agregada gracias a manejadores de bases de datos, tablas, índices, buscadores, etcétera.

Con el advenimiento de los libros electrónicos, el tema de los formatos para textos vuelve a estar en boga para análisis más detallados. Originalmente se comenzó por editar prácticamente todo en formato “PDF”. Desde el punto de vista técnico ofrece muchas ventajas las cuales lo habían estado convirtiendo en el estándar *de facto* para ese propósito. El inconveniente de este formato es que no puede ser encriptado

o bloqueado efectivamente en lo relativo a copias. Por ello, en lo tocante a ciertos libros electrónicos cuyo propósito sea la distribución masiva y no importa mucho el copiado, sigue siendo una excelente opción; el problema se presenta cuando el propósito del libro es la venta y donde el hecho de que no pueda ser reproducido libremente se vuelve factor indispensable. Otras empresas han comenzado ya a liberar sus formatos para este tipo de material; entre ellas distinguimos el “OEB” de la “Electronic Book Federation” y el formato “LIT” de la compañía “Microsoft”, pero éstos todavía no se han consolidado como estándares por una serie de detalles que analizaremos posteriormente en un apartado especial para libros electrónicos.

El Formato “OpenDocument” —*Documento Abierto para Aplicaciones de Oficina Automática*— es una propuesta de formato estándar para archivos de documentos “*ofimáticos*” tales como oficios, memorandos, hojas de cálculo, gráficas y presentaciones. Su desarrollo ha sido encargado a la organización OASIS y está basado en un esquema XML inicialmente creado por la organización *OpenOffice*. El estándar fue desarrollado públicamente por un grupo de importantes organizaciones del cómputo a nivel mundial, entre ellas algunas como IBM, Sun, Novell, Intel, Adobe, etcétera. Es accesible públicamente, y puede ser implementado por cualquier persona o grupo sin restricciones. Este formato pretende ofrecer una alternativa abierta a los formatos de documentos propietarios de Microsoft tales como .doc, .xls, .ppt, e incluyen el denominado “*Microsoft Office Open XML Format*” cuyos requisitos de licenciamiento le impiden su empleo a diversos competidores. La motivación principal para usar formatos estándar estriba en que las personas y organizaciones que lo hacen, evitan convertirse en clientes cautivos de un único proveedor de software, lo que les permite cambiar de ambiente informático si su proveedor actual es expulsado del mercado o cambia su modelo de licenciamiento en términos menos favorables para el cliente [Oasis Standards..., 2005].²⁷

²⁷ *Oasis Standards and Other Approved Work*. 2005. Página Oficial de la Organización OASIS. Disponible noviembre 3, 2005 en: <http://www.oasis-open.org/specs/index.php>

OpenDocument es el único estándar para documentos ofimáticos implementado por distintos competidores, aceptado por organismos de estandarización independientes y susceptible de ser implementado por cualquier proveedor. Ya existen versiones de programas que usan el formato *OpenDocument*, siendo el más conocido *KOffice*, a partir de su versión 1.4, liberada en junio de 2005—, así como la versión 2.0 del conocido *OpenOffice* la que lo utiliza como formato de archivos por omisión.

Mención aparte merece el lenguaje “XML” —*Extended Markup Language*—, lenguaje extendido de marcado, el cual, dadas sus características, parece perfilarse como el heredero de HTML y PDF. Por su importancia, merece tratarse con detalle. Se discutirá más adelante en un apartado especial sobre lenguajes de marcado.

Cabe mencionar en este punto que un formato no es patentable; no se deben pagar derechos por su uso. Por tanto, cualquier biblioteca o persona en teoría pueden generar documentos en un formato dado sin invertir en el uso del formato. Lo que se paga es el uso de un programa o dispositivo que genere o interprete los documentos del formato en cuestión en una computadora, pero no en sí por usarlo. Por ejemplo, la compañía *Adobe* vende su programa que permite crear documentos en su formato “PDF”, pero regala la parte que permite leerlo. De esta forma, una biblioteca que crea documentos en ese formato puede instruir a sus lectores para que obtengan el programa lector de ese formato en forma gratuita y poder acceder así a esa información digital. La biblioteca puede pagar el programa que genera los documentos. Más aún, si la biblioteca puede generar esos documentos en ese formato sin necesidad de usar ese programa, puede construir su colección en esa forma sin pago alguno por el formato, como por ejemplo los documentos “.PDF” generados con “*OpenOffice*”.

2.4.2 Imagen fija

Se han discutido ya algunas características de los formatos especiales para el manejo de textos electrónicos. Si bien los conceptos básicos

son los mismos en lo referente a imágenes, es conveniente profundizar en ellos para comprender mejor sus características.

Como ya se mencionó al tratar textos, el proceso de digitalizar una imagen consiste básicamente en el registro de cambios en intensidad de luz y color reflejados por un objeto, o asignados al mismo. Si pensamos en ese objeto como una imagen plana —impresión fotográfica, grabado, mapa, etcétera— podemos pensar en ese documento como una “matriz” o arreglo de puntos distribuidos a lo largo y ancho de un plano; también conocidos como *bitmap* o “mapa de bits”. Básicamente existen dos técnicas para lograr esto; la primera es conocida como *rastering*: un escáner por ejemplo, “romperá” esa imagen en miles y miles de pequeños puntos, asignándole a cada uno de ellos un valor numérico para el color y posición que ocupa en el plano, creando así el mapa de bits. El valor numérico del color se obtiene de una tabla preestablecida; es fácil imaginar que si una imagen es sólo blanco y negro, la paleta es únicamente de dos colores. Eso está bien para diagramas y tablas con líneas y trazos; pero en la realidad una imagen blanco y negro tiene tonos de gris. Una imagen típica de un escáner en blanco y negro puede tener como paleta entre 4 y 16 tonos de gris. Cuanto mayor sea esa paleta, mejor fidelidad cromática. Si la imagen es a colores, se pueden preestablecer paletas de 24, 32, 256, 65,535 (64 x 64), 16 millones de colores, etcétera. Un buen escáner puede preestablecer hoy en día más de 64 millones de colores distintos, mucho más allá de lo que el ojo humano puede diferenciar. Como puede verse, por fidelidad cromática no hay impedimento; el problema estriba en que, a mayor paleta cromática, el valor numérico del color de cada punto dado es un número muy grande, el cual representado en binario ocupa muchos “bits”; es decir, mucho espacio de almacenamiento.

Para explicar esto, conviene introducir la otra característica de una imagen bajo esta técnica, su “resolución”; es decir, el número de puntos en que se “rompe” la imagen a ser digitalizada. A estos puntos se les denomina “píxeles”. Puntos grandes —*baja resolución*— hacen una imagen cuyos bordes y trazos pueden ser observados a simple vista, y por tanto burda y de poca calidad. Entre más finos sean los puntos —*alta resolución*— la imagen mejora su definición y por ende su ca-

lidad. La resolución se mide en “puntos por pulgada”, “ppp” o por sus siglas en inglés, “dpi”, *dots per inch*. Así, al tener una imagen original que mida por ejemplo, cinco por tres pulgadas, al seleccionar una resolución para digitalizar de, digamos 300 dpi, significa que el escáner “romperá” esa imagen en 300 puntos por cada pulgada, y esto dará $300 \times 5 = 1,500$ puntos horizontalmente, y $300 \times 3 = 900$ puntos verticalmente; dicho de otra forma, cada pulgada cuadrada de la imagen se rompe en $300 \times 300 = 90,000$ puntos. Esto es, un total de 1’350,000 puntos para representar el total de esa imagen. En otro ejemplo, una resolución seleccionada de 75 dpi indicaría que esa imagen posee $75 \times 5 = 375$ puntos horizontalmente y $75 \times 3 = 225$ puntos verticalmente; esto es, 5625 puntos por cada pulgada cuadrada, y 84,375 puntos para el total de la imagen. Al observar las dos imágenes resultantes, puede percibirse que la primera tiene mejor definición y nitidez, puesto que está formada por puntos más finos; la segunda, digitalizada a 75 dpi es una imagen más o menos burda y de poca resolución o nitidez, pero la diferencia en espacio ocupado es enorme, y no en un factor de 4, la relación entre 75 y 300 dpi. La diferencia en espacio ocupado es de 16 veces, el cuadrado de 4. Los escáners actuales trabajan típicamente estas resoluciones y llegan fácilmente hoy día hasta 2400 dpi y todavía más, o sea imágenes ultra finas. Y siguen mejorando cada día, pero el espacio ocupado por cada imagen es realmente enorme.

Si cada pixel es representado por un número binario o “bit”, es decir, un uno o un cero, el archivo de la primera imagen mediría 1’350,000 bits; el de la segunda 84,375 bits. Pero un bit uno o cero sólo puede representar dos colores, típicamente blanco y negro. Si deseamos que los puntos sean de color, una secuencia de bits debe ser seleccionada para representar el valor numérico del color dentro de una paleta. Una paleta de grises con 16 colores requerirá de pixeles de 4 bits ($16=2$ a la cuarta potencia); una paleta de 256 colores requerirá de 8 bits por cada pixel ($256=2$ a la octava potencia) y una de 65,536 colores tendrá pixeles de 16 bits cada uno ($65,536=2$ a la decimosexta potencia). Dado que un “byte” o unidad de medida de almacenamiento contiene 8 bits, al multiplicar el número de pixeles de una imagen por el número de bites requeridos para cada pixel, el número de bytes de una imagen

puede llegar fácilmente a varios cientos de miles o millones de bytes por cada imagen.

Entonces la resolución y la riqueza cromática son un arma de dos filos: una imagen de alta resolución proveniente de una hoja, digamos tamaño carta, está formada por miles y miles de puntos, cada uno de los cuales a su vez requiere de un número más o menos grande de bits cada uno dependiendo de su paleta cromática. Puede verse entonces que una imagen carta de alta resolución y buena paleta tiene millones de puntos con valores numéricos grandes cada uno. Esto implica un gran número de “bits” consumidos para esa imagen digitalizada; transmitirla por la red tomará más tiempo. Almacenarla requiere de un gran espacio en un disco duro u otro dispositivo similar. Si la cantidad de imágenes que se van a digitalizar es grande, el espacio requerido para su almacenamiento puede llegar a ser enorme, lo cual conlleva un costo oneroso. Si bien existen técnicas para compresión de estos archivos con objeto de que ocupen menos espacio en la vida real, a las cuales me referiré más adelante, la compresión puede implicar pérdida de calidad de la imagen.

Existe entonces una importante relación de proporción y balance en este sentido: a menor riqueza cromática y a menor resolución, menor calidad final de la imagen, pero es más económica en el espacio ocupado. A mayor color y mayor resolución, la imagen resultante tiene mayor calidad, pero ocupa más espacio y cuesta más. ¿Qué criterio debe prevalecer entonces, calidad o economía?

Por esta razón y para poder responder, *el propósito de ese proceso de digitalización se vuelve muy importante*. Como ya se mencionó, las principales razones para digitalizar son preservar o distribuir. De ellas puede extraerse ese criterio necesario para tomar la decisión. Si nuestro propósito en ese momento es preservar un documento único o raro y por ello muy valioso, y queremos obtener un nuevo “original” de él para preservarlo, debe hacerse el gasto. Debe digitalizarse con la mayor calidad posible poniendo en segundo plano el costo. Al fin y al cabo, los documentos raros o únicos no se dan por millones. Entre más valioso y escaso sea un documento, mayor calidad debe imprimirse en el proceso de digitalización.

Por el contrario, si el propósito principal es digitalizar para distribuir, para poner al alcance de los usuarios alguna colección, el criterio principal es la economía, sobre todo si el documento no se considera raro o escaso. Por ejemplo, si se habla de digitalizar un incunable para preservarlo, debe hacerse con el criterio de la mayor calidad, al margen de los costos mayores. Los incunables no se dan en maceta. Si se habla de digitalizar los “apuntes” de apoyo a las asignaturas de una escuela, deberá prevalecer el criterio de la economía.

Por supuesto, el criterio de la economía no debe llegar a niveles que reduzcan tanto la calidad que hagan el material ilegible o no apto para su uso. Debe haber una calidad “mínima” con la cual el material le es útil al usuario y no obstante sigue resultando económico de almacenar. Cuando se saca una fotocopia de un artículo de una revista en una biblioteca para un usuario, estamos conscientes de que esa copia no es de la calidad del original, pero si es suficientemente nítida, le servirá a aquél perfectamente para sus propósitos. Sólo en ciertos casos requerirá él una fotocopia a color, o algo parecido. Si la fotocopidora tiene una calidad razonable, no hay problema. Si la máquina se descompone y entrega copias ilegibles, ese material es inútil. Es por ello muy importante definir cuál es la “calidad suficiente” que un cierto material digitalizado debe tener para ser distribuido, lo cual reduciría a la vez el espacio requerido y con ello, los costos. Entonces, el concepto de *reproducción digital fiel* es un término relativo a la funcionalidad y valor de uso de un documento, lo cual entonces, define su aceptabilidad. Las copias de “alta calidad” para preservación pueden almacenarse a su vez en dispositivos más baratos, como DVD, CD-ROM, cintas o cartuchos; es decir, almacenamiento que no esté en-línea. Puede obtenerse a su vez de ellas copias de menor calidad y costo con fines de distribución, y que son las que podrían estar en-línea al servicio del público. Como resumen de lo anterior, he aquí una pequeña tabla con los tamaños de archivo que resultan para una imagen tomada con diferentes valores de resolución, paleta y compresión: [Morrison, 1998].

Resolución (dpi)	400 dpi	300 dpi	200 dpi	100 dpi
2-bit blanco y negro	20Kb	11Kb	5Kb	1Kb
8-bit grises o color	158Kb	89Kb	39Kb	9Kb
24-bit color	475Kb	267Kb	118Kb	29Kb

300 dpi 24-bit, imagen a color: 2.65 x 3.14 pulgadas:
TIFF sin compresión: 2188 Kb. JPEG con mínima pérdida: 59 Kb

300 dpi 8-bit, imagen a color: 2.65 x 3.14 pulgadas:
TIFF sin compresión: 729 Kb. JPEG con mínima pérdida: 76 Kb

100 dpi 24-bit, imagen a color: 2.65 x 3.14 pulgadas:
TIFF sin compresión: 249 Kb. JPEG con mínima pérdida: 9 Kb

100 dpi 8-bit, imagen a color: 2.65 x 3.14 pulgadas:
TIFF sin compresión: 85 Kb. JPEG con mínima pérdida: 12 Kb

Con las cámaras digitales como medio para digitalizar imágenes sucede algo parecido, ya que se basan en el mismo principio: pueden preseleccionar paleta cromática casi con la misma riqueza que un escáner; pueden tomar imágenes con mayor o menor resolución y “romper” la imagen en un número mayor o menor de puntos, etcétera. La mayor diferencia estriba en que la cámara no trabaja con una unidad de “pulgadas cuadradas”, *dpi*, como el escáner. La “ventana” de una cámara digital tiene un visor que es siempre igual, típicamente rectangular en una proporción de 5 a 4 en relación de la distancia horizontal con la vertical. En esa “ventana” se alojarán un cierto número de puntos que forman una imagen para una fotografía dada gracias a un “Dispositivo de Carga Acoplada” o CCD *Charge-Coupled Device*. Éste es un dispositivo sensible a la luz, en forma de rectángulo, formado por miles y miles de celdillas que generan electricidad al recibir la luz. En comparación, sería el “rollo” de película de una cámara fotográfica común. Cada celdilla del CCD está recubierta con un filtro sensible a las tonalidades azul, verde y rojo. Al número total de puntos de celdillas en esa ventana que es siempre de la misma superficie, generalmente millones de puntos, se le conoce como *megapíxeles* en vez de *dpi*. Un megapixel es poco más de un millón de puntos.

La diferencia entonces consiste en que el escáner refiere sus unidades a puntos por pulgada cuadrada de imagen tomada, multiplicados por el número total de pulgadas cuadradas que conforman esa imagen. En la cámara digital la superficie siempre es la misma, por lo que las unidades refieren el número total de puntos que conforman esa imagen.

Así, si una cámara “rompe” una imagen en su ventana en, digamos 1280 por 1024 puntos, el total de la imagen fotografiada es de 1’310,720 puntos; en números redondos 1.3 megapíxeles. Una cámara de 5.2 megapíxeles máximos indica que puede tomar varias resoluciones, hasta un máximo de 5.2 millones de puntos totales por cada imagen. Esto es, imágenes de 2560 x 2048 puntos. En la actualidad hay cámaras de hasta 21 megapíxeles totales por imagen, lo cual equivale a una resolución de 5120 x 4096 puntos, sumamente alta, y siguen mejorando; como referencia, el ojo humano tiene alrededor de 127 millones de células cromáticas o 127 megapíxeles. Puede pedirse a la cámara que use menor resolución respecto a su máximo, pero por supuesto, no pueden tomarse imágenes más allá de su resolución máxima estipulada. Debe tenerse en mente que mayor resolución en la fotografía implica que pueden tomarse menos fotografías con una cierta dotación fija de memoria interna, su “rollo” digital por así decirlo, y viceversa. Una cámara con 4 megabytes de memoria puede obtener alrededor de 100 fotos de baja resolución con esa memoria, o 25 de resolución media, u 8 de alta resolución. La ventaja de la cámara es sin duda su portabilidad mucho mayor que la de un escáner. En la toma de imágenes de exteriores es sin duda una gran ventaja, o cuando el documento no puede ser desencuadrado, como en el caso de un libro antiguo.

La segunda técnica para representar imágenes en un archivo se conoce como *rendering* o de “Meta/Vectores”. Esta técnica tiene su origen y es ampliamente utilizada por los programas que permiten “dibujar” o crear imágenes en la computadora. En este tipo de representación una imagen no es más una matriz de puntos contiguos o “mapa de bits”; consiste en vectores de líneas y colores almacenados en fórmulas matemáticas. Esto permite las acciones conocidas hoy en estos programas para cambiar escalas de imágenes, girarlas, realizar efectos cromáticos complejos, etcétera, típicas de un programa de este tipo. También

pueden contener información embebida de datos cromáticos, los “metadatos” de la imagen, o mezclas de metadatos y vectores, de ahí el nombre.

Prácticamente todos los formatos utilizados hoy en día para la digitalización de imágenes obedecen a los principios y técnicas enumerados anteriormente [Harvard Library Preservation, 2000]. Entre los formatos más utilizados en la actualidad con fines de preservación, encontramos el formato “TIF” o “TIFF” de “Adobe”, así como el formato “TIFF EP”, normas ISO TC/42, ISO 12234-2; el formato “PCD” —Photo Compact Disc Image graphics file— de “Eastman Kodak”, y el estándar “BMP”, “bitmap”. Todos estos han sido creados para máxima calidad, resolución y croma. De más está decir que consumen mucho espacio.

“TIFF” —Tagged Image File Format— es un formato muy flexible y extensible que es soportado ya por una amplia variedad de aplicaciones de procesamiento de imágenes y plataformas. Entre sus principales características se encuentran: el soportar varios tamaños de pixel y por tanto paleta variable; un archivo puede contener varias imágenes; puede usar distintos algoritmos de compresión, un sólido conjunto de metadatos embebido, etcétera. Dada la alta calidad de imágenes que produce y que es el estándar de los equipos “Apple”, diseñados con enormes ventajas específicas para la creación y manejo de imágenes, se ha convertido en un estándar ampliamente aceptado para la preservación de imágenes.

Entre los más utilizados para distribución de imágenes electrónicamente; esto, es vía cd-rom, *Web*, etcétera, encontramos los tres siguientes formatos:

“GIF” —Graphic Interchange Format— el cual fue desarrollado originalmente por la empresa “Compuserve”, y cuya patente pertenece hoy a la compañía “Unisys”. “GIF” ha sido durante mucho tiempo uno de los formatos de imagen más populares sobre la Internet, sobre todo debido a su tamaño compacto. Es ideal para pequeños íconos de navegación, diagramas simples e ilustraciones, es decir, ahí donde se requiere precisión, o gráficas con bloques grandes de un solo color. El formato es “sin pérdida”, lo que significa que no se hace borroso o sucio. Los archivos “GIF” usan una técnica simple llamada compresión

LZW para reducir el tamaño de las imágenes al encontrar modelos repetidos de píxeles, sin que ésta compresión degrade la calidad de imagen. Puede manejar una paleta de hasta 256 colores.

El formato “JPG” o “JPEG”—Joint Photographic Expert Group— fue diseñado por esa asociación con fines de distribuir fotografía digital.²⁸ Es capaz de desplegar millones de colores en su paleta y maneja perfectamente la mezcla compleja en cuanto a matices que ocurre en imágenes fotográficas. Como tiene una capacidad muy grande de compresión se pueden obtener archivos muy pequeños; por esta razón, aunado al hecho de que su uso es “libre de regalías” se ha difundido enormemente para uso en páginas *Web*, dado el efecto “aligerador” sobre ellas; fue de hecho un estándar *de facto* para distribución de imágenes en la red antes de volverse un estándar de norma (ISO/IEC 10918 – 1:1993(E)). Este formato usa un algoritmo de compresión complejo que puede ser aplicado sobre una escala móvil; la compresión es conseguida “olvidando” ciertos detalles sobre la imagen, que el formato tratará entonces de llenar más tarde cuando esté siendo mostrada. Puede salvarse una imagen “JPG” con la compresión del 0 % para una imagen perfecta con tamaño de archivo grande, o con compresión del 80 % para una imagen pequeña pero perceptiblemente degradada. En el uso práctico, un ajuste de compresión aproximadamente del 40 al 60 % llevará al equilibrio óptimo entre calidad y tamaño con fines de distribución en *Web*. El formato JPEG actual muy probablemente será sustituido en algún punto en el futuro por el formato “JPEG2000” actualizado.²⁹

El “PNG”—*Portable network Graphics*— es un formato recién “inventado” expresamente para la *Web* en respuesta al esquema de licencias introducido por “Compuserve” que implicó para los creadores de cualquier programa que soportara su formato “GIF” el pago de cinco mil dólares por ese derecho; por lo mismo, el uso de PNG está libre de

28 Joint Photographic Experts Group. 2003. Página *Web* Oficial del Grupo de Expertos. Disponible noviembre 3, 2005 en: <http://www.jpeg.org>

29 JPEG 2000. ISO/IEC 15444-1:2000, “*Information Technology — JPEG 2000 Image Coding System — Part 1: Core Coding System*”. July 31, 2002, Final Committee Draft (FCD). Disponible noviembre 3, 2005 en: <http://www.jpeg.org/public/fcd15444-1.pdf>

regalías y actualmente está avalado por el “Consortio Mundial para el desarrollo de la World Wide Web”, conocido también como w3c. Una versión del formato, PNG-8, es similar al formato “GIF” en sus funciones y también maneja una paleta de 256 colores, pero con mayor eficiencia en la compresión y por tanto en el tamaño del archivo resultante. El PNG-24 es otra variante que maneja una paleta de colores enorme, semejante al “JPG”, aunque sin la eficiencia de compresión de aquél. La característica distinta de PNG sobre los otros es su manejo de imágenes desde “totalmente transparentes” hasta “totalmente opacas” lo cual permite, de manera interesante, sobreponer imágenes, marcas de agua, tapices, etcétera, en las páginas *Web*. Es muy pronto todavía para decir si substituirá a “GIF” y a “JPG” en la edición en *Web*, ya que no todos los navegadores pueden desplegarlo actualmente. [World Wide Web Consortium, PNG. 2002].

El formato “PIC” de *Macintosh Quickdraw* es sumamente utilizado en la industria editorial para formar los originales de libros y revistas. Este formato comprime muy eficazmente las imágenes con tonos parecidos sin perder calidad. Trabaja con 16 o 32 bits de color.

Finalmente cabe mencionar que existen muchos otros formatos para almacenamiento de imágenes fijas: tga, pcx, clp, dcx, fpx, img, mac, msp, pct, ppm, psd, rle, raw, wpg, rif, etcétera, son sólo algunos de los acrónimos que representan formatos para almacenamiento de imágenes. Con el advenimiento de los programas de computadora conocidos como “editores gráficos”, pero sobre todo con las cámaras digitales, toda una colección de ellos han aparecido. La presentada en los párrafos anteriores no fue una lista exhaustiva; simplemente una reseña de lo más aceptado en el momento en el mundo de las imágenes dentro del ambiente de las bibliotecas digitales.

2.4.3 Audio

Para analizar la digitalización de pistas de audio, es necesario establecer primero los soportes anteriores de este tipo de documentos. En esencia, el sonido es una forma de onda de presión que se desplaza en el aire; por ser una onda, es una señal de variación continua: cambia

de instante en instante, y mientras cambia entre dos valores dados, pasa por todos los valores intermedios. Los sonidos han sido grabados por casi un siglo en dispositivos “analógicos” antes de la aparición del disco compacto CD. Un dispositivo analógico funciona, de ahí su nombre, por una “analogía” o simulación del fenómeno real. En un disco de acetato o vinilo, la analogía se logra por medio de la profundidad y depresiones de los surcos en el disco con respecto a las ondas de presión; la aguja, al recorrer el surco, recoge esas formas y por analogía las convierte en un impulso eléctrico que es amplificado por un circuito y pasa a las bocinas. De modo semejante funciona un casete o cinta magnetofónica: el magnetismo grabado ahí tiene analogía con la forma de la onda de sonido; al ser “leído” por una cabeza magnética, es transformado nuevamente por analogía a un impulso eléctrico usado por el amplificador y las bocinas para producir de nuevo el sonido. El proceso de “digitalización” del sonido, ya sea voz o música, se logra grabando los tonos, volúmenes, pausas e intensidades del sonido en un formato digital; esto es, en una representación numérica convencional, como ya hemos comentado, tomando “muestras” de la onda a intervalos de tiempo iguales y preestablecidos. El valor de la muestra es un número igual a la amplitud de la señal en el instante del muestreo.

El principio de la digitalización que estudiamos en imágenes fijas se mantiene: existe el equivalente de la “resolución” en el sonido. Se mide en “tasa de bits” —*bits rate*— por segundo o algún equivalente; ésta es la tasa de muestras a obtener por unidad de tiempo. A mayor tasa de bits, mayor cantidad de información para los circuitos, canales y bocinas y el sonido tendrá por lo mismo mayor calidad; pero conlleva a la vez mayor espacio consumido para almacenamiento y en consecuencia mayor costo. A menor tasa de bits la cantidad de información que representa al sonido disminuye y con ello decrece su calidad, pero también el espacio requerido. Los formatos creados y utilizados al efecto siguen estos principios básicos.

Un trozo de audio es grabado entonces en un archivo de computadora en forma numérica bajo un cierto patrón o “formato”. Típicamente un formato de audio contiene tres partes: en primer lugar la “envoltura” —*wrapper*—, en la cual se registran características accesorias del archivo: licenciamiento, capacidad de reproducción en-línea del sonido

—*streaming capability*—, la “letra” o texto de la pieza musical, etcétera. En segundo lugar, un “encabezado” —*header*—, el cual almacena información acerca del archivo en sí, tales como la resolución, tasa de muestreo del sonido, tipo de compresión, etcétera. Finalmente el archivo contiene los datos de información propios del sonido de la pieza de audio, bajo una cierta estructura o tecnología típica para la codificación y decodificación del mismo, conocida como “*codec*”. Entre los tipos de codificación o “*codecs*” más utilizados se encuentra el “PCM” —Pulse Code Modulation— o *código de muestreo de la señal de audio*. Dado que es un código genérico, puede ser leído por la mayoría de los sistemas de audio, de manera similar a la que un archivo de texto plano tipo .TXT puede ser leído por casi todos los programas de procesamiento de texto. PCM es utilizado como base para los reproductores de discos compactos musicales y cintas digitales de audio “DAT” o “Digital Audio tapes”. PCM es también el “codec” que usan sistemas de computadora muy comunes como “AIFF” —Audio Interchange File Format— y “WAV” —Wave Form Format—.

Los formatos más utilizados para preservación o almacenamiento de largo plazo para documentos de tipo audio tienen por lo general altas tasas de muestreo del sonido y poca o nula compresión. La calidad del sonido en estos formatos es por lo tanto alta pero consumen por lo mismo grandes cantidades de espacio en el archivo y en su consecuente almacenamiento. Conviene en este momento hacer una lista de los formatos de audio más utilizados en esta época:

- CDDA (Compact Disc Digital Audio). Es el formato nativo de los discos compactos, con frecuencias de muestreo de 44.1 KHz, 16 bits de cuantificación y dos canales. Tiene alta calidad, pero por lo mismo ocupa mucho espacio. Un disco compacto musical promedio tiene capacidad cuando mucho para 74 minutos de grabación.
- WAV (Wave-Form). Uno de los estándares en las computadoras con sistema operativo *Windows*. A diferencia del CDDA, posee más frecuencias de muestreo: desde los 8 khz hasta 192 khz, por lo que puede ser utilizado para archivos muy comprimidos y de poca calidad, hasta archivos con muy alta calidad, aún mayor

que la del CDDA. Muy utilizado en bibliotecas, no se ha posicionado como estándar universal.

- AIFF (Audio Interchange Format File). Típico de las computadoras marca “Macintosh” de “Apple Computers”. Soporta hasta 44.1 khz y 32 bits de cuantificación. Muy utilizado para la preservación de archivos de audio en bibliotecas.
- RA o RM (Real Audio). Creado por la empresa *Real Networks*. Puede lograr muy buena compresión y descompresión, pero por lo general su calidad no es adecuada para aplicaciones profesionales; es muy aceptado en la distribución de señales de audio a través de Internet *en vivo* —llamado “*en tiempo real*”, o “*streaming*”—.
- Windows Media Audio (WMA o MS Audio). Creado por Microsoft para competir con el MP3. Aún está en duda su permanencia y aceptación.
- MIDI (Musical Instrument Digital Interface). El formato en sí es un protocolo o norma para transferir información entre dispositivos musicales como sintetizadores, guitarras eléctricas, violines electrónicos y software diverso. Un archivo MIDI indica tonos, pulsos, variaciones de intensidad, sostenidos y duraciones, que pueden modificarse en cada instrumento. Es muy útil para los artistas que desean generar sus producciones combinando varios dispositivos con el apoyo de computadoras. Es muy conocido también por ser el que se utiliza para la música de juegos de computadoras, tonos de celulares, etcétera. En la práctica no es de alta calidad.

En cuanto a las empresas que los producen o impulsan, AIFF ha sido el formato estándar para las computadoras “Macintosh” de “Apple Computers”, mientras que “WAV” o “Wave” ha sido el formato estándar para audio digital en computadoras personales con sistema operativo “Windows”. Recientemente esta empresa ha introducido un nuevo estándar de formato de audio para su sistema denominado “WMA”, —Windows Media Audio—, al tiempo que Apple ha introducido el formato “QT” —QuickTime— como su nuevo estándar. Otro formato muy difundido

en este tipo de archivos es el denominado “RA” o “RAM” —Real Audio—, lanzado por la empresa “Real Networks”; la empresa “Sun Microsystems” ha establecido el formato “AU” como el estándar para sus máquinas. Casi todos ellos garantizan alta calidad pero consumen mucho espacio, un promedio de diez Megabytes por minuto de grabación.

Por el lado de los formatos para distribución, de menor calidad pero mucho menor consumo de espacio, el estándar indiscutible del mercado, si bien *de facto*, es el formato denominado “MPEG-1 Audio Layer-III” o “MPEG-1 Capa 3”, más conocido como “MP3”. Fue creado por *Thomson Multimedia* y el Instituto *Fraunhofer* de acuerdo con la norma establecida por el Grupo de Expertos de Imágenes en Movimiento —*Moving Picture Experts Group* o MPEG— para ser el estándar de audio digital en los archivos de video MPEG-1. Es ampliamente utilizado a nivel mundial para intercambio de audio por la red, ya que puede obtener factores de compresión de hasta doce veces con respecto a los enunciados anteriormente para preservación; ello se basa en el principio de eliminar las frecuencias que no son audibles por el oído humano. Por lo mismo, el formato tiene varias calidades de salida según la compresión utilizada, pero en promedio puede consumir alrededor de un Megabyte por minuto de grabación con una calidad aceptable, con lo que obtiene de diez a doce veces más rendimiento en los archivos respecto a los formatos de preservación; esto lo hace *muy aceptado en las bibliotecas con fines de distribución* de información tipo audio. Su sucesor, quien también maneja video, parece ser alguno de entre el “MPEG-1 capa 4” o “MPEG AAC” o el “MPEG-2”. Estos formatos, junto con el “MIDI” —Musical Instrument Digital Interface—, son ampliamente utilizados en aplicaciones de distribución de audio a través de la red dado su mínimo consumo de espacio.

Éstos no son los únicos formatos para audio digital, pero son los más utilizados alrededor del mundo de las bibliotecas, las redes y la información. Una lista exhaustiva puede estudiarse en los sitios *web* de la “Audio Engineering Society”, <http://www.aes.org> y del “Internet Sound Institute” <http://www.soundinstitute.com>. Finalmente, a manera de resumen presento una tabla de lo enunciado anteriormente:

Tipo	Extensión(es)	Codec
AIFF (Mac)	.aif, aiff	*PCM
AU (Sun/Next)	.au	*u-law
CD-audio (CDDA)	No tiene	PCM
MP3	.mp3	MPEG-1 Audio Layer-III
Windows Media Audio	.wma	Propietario (Microsoft)
QuickTime	.qt	Propietario (Apple Computer)
Real Audio	.ra, .ram	Propietario (Real Networks)
Wave Form	.wav	*PCM

*Puede ser usado con otros *codecs*.

(Fuente: <http://www.teamcombooks.com/mp3handbook/12.htm>)

FIGURA 2.2 TABLA DE FORMATOS DE AUDIO DIGITAL MÁS COMUNES

2.4.4 Imagen dinámica

Con respecto a la imagen en movimiento, video o cine, encontramos los mismos principios que rigen los anteriores: formatos para preservación, de alto consumo de espacio, y formatos para distribución, de menor calidad pero más económicos al almacenar. En este caso, el problema se agrava por el hecho de que la imagen en movimiento, como es sabido desde hace mucho gracias a la tecnología del cine, requiere de 24 imágenes por segundo de grabación; esto hace que el video sea tremendamente consumidor de espacio. Puede ser reducido o comprimido a 18 o 16 cuadros, pero a costa de la calidad, y por ello deben usarse cuadros pequeños en la pantalla.

El problema con el mundo del video se agudiza dada la enorme variedad de formatos, tanto analógicos como digitales, que se utilizan hoy en día. Existen formatos para transmitir, para almacenar, para vender o alquilar. Se usan diferentes estándares en América y en Europa; video casero, etcétera.

Distingamos primero entre los formatos estándares de la industria del video analógico; entre éstos se encuentran tres formatos principales a nivel mundial: NTSC, PAL y SECAM;³⁰ en Norteamérica usamos el primero. Éstos son formatos de producción y transmisión del video analógico por parte de la industria televisiva. A su vez, éstos tienen que ser transformados a otro para su almacenamiento y distribución; entre ellos conocemos los formatos Beta, VHS, DVD, VCD, etcétera, gracias a la industria del video. Con fines de preservación o distribución se han usado diferentes combinaciones de diversos grupos de estos formatos; por ejemplo, una película se almacena con NTSC y DVD; es decir, formato americano analógico convertido a un disco DVD digital.

La industria televisiva actual ha creado ya formatos digitales, con miras a mejorar su calidad intrínseca. Todos de gran consumo de bytes por minuto de grabación. Centrándonos en nuestro objeto, que son los documentos almacenables y distribuibles en bibliotecas digitales con ayuda de una computadora, seleccionaremos los más utilizados a este respecto en la actualidad.

El formato “QuickTime” fue desarrollado por la empresa “Apple”, la cual, como ya se ha mencionado, ha diseñado sus equipos con una fuerte especialización para el manejo de imágenes; puede ser utilizado también en equipos personales —PC— con un programa manejador. Este formato soporta los mecanismos de codificación de video digital más utilizados hoy en día, tales como “Cinepak”, “JPEG” y “MPEG”. En 1988 la oficina ISO lo seleccionó como la base para su nuevo estándar “MPEG-4”.

El formato “AVI” —*Audio Video Interleaved*— o “audiovideo entrelazado” es el formato general de la empresa “Microsoft” para máquinas de tipo PC bajo su sistema “Windows”, y por ello se ha convertido en un estándar *de facto* en el medio. AVI ha desembocado en el más extenso formato multimedia de Microsoft denominado “Windows Media 9”.

Con fines de preservación de video, debe tenerse cuidado especial en la actualización de formatos de este tipo. Como ya se ha mencionado,

30 ISO/IEC 15444-2:2004, “*Information Technology —JPEG 2000 Image Coding System— Part 2: Extensions*”, May 15, 2004, Final Committee Draft (FCD). Disponible noviembre 3, 2005 en: <http://www.jpeg.org/public/fcd15444-2.pdf>

más de cincuenta de ellos han existido desde 1956, cuando se empezó a grabar el video. De hecho, puede observarse un catálogo completo de formatos de video en el sitio de Paul Messier, del grupo “Boston Art Conservation” [Messier, 2003] divididos en seis categorías en función de su obsolescencia, desde “extintos” hasta “vigentes o de bajo riesgo”. Como resumen, he aquí una lista de lo más consensado en cuanto a formatos digitales de video para ser utilizados en la red:

- *En baja resolución*
Real Media 56 kpbs (kilobytes por segundo)
Windows media 28 kpbs
OT QuickTime 56 kpbs
- *En resolución intermedia*
Real Media 256 kpbs
Windows Media 384 kpbs
OT QuickTime 256 kpbs
- *En alta resolución*
Mpeg-1 1.5 mbps (megabytes por segundo)
Mpeg-2 2 mbps

Nuevamente, al igual que en las secciones anteriores, no es ésta una lista exhaustiva de formatos de video sino sólo una visión de lo más conocido y reconocido en el medio de las bibliotecas digitales.

2.5 COMPRESIÓN Y ALMACENAMIENTO

Como ha podido observarse, en todos los formatos electrónicos utilizados en proyectos de bibliotecas digitales surge siempre una interrogante para el bibliotecario: el balance adecuado entre calidad y consumo de espacio. A mayor calidad del documento digitalizado crece el espacio requerido para guardar el material en una forma no-lineal; y mayor espacio implica mayor costo de almacenamiento y mayor tiempo de transmisión del documento. Ello marca una separación definitiva, por lo menos para la próxima década, de los formatos utilizados para preservar y distribuir material por parte de las bibliotecas. La

confusión entre las características y costos de unos y otros marca muchas veces que un proyecto de digitalización se realice o no; por ello es necesario que los bibliotecarios estén conscientes de esta característica de un documento digital, aunque sea de manera general, y también de sus costos asociados.

Hemos observado que tamaño y costo de un documento son factores totalmente entrelazados y directamente proporcionales. Por ello se invierten en la actualidad muchos esfuerzos en investigación buscando mejorar esos rendimientos y “comprimir” más y más los documentos. Gran parte de los formatos que han aparecido recientemente son el fruto de esas investigaciones.

En efecto, el advenimiento de los documentos multimedia y de la transmisión masiva de información en Internet ha creado un auge inusitado en la investigación y el desarrollo de las técnicas de compresión de la información, ya que la cantidad de “bytes” o caracteres de una pieza de información cualquiera incide sensiblemente en el tiempo en que un dispositivo tarda en enviarla o recibirla; sin el desarrollo de estas técnicas, los teléfonos celulares no serían prácticos, ni tampoco el audio o video transmitido entre computadoras. Por ejemplo, la representación digital de un segundo de video sin compresión, usando el formato CCIR 601, ocupa más de 20 Megabytes, es decir el equivalente 20 millones de caracteres. Si consideramos el número de segundos de una película promedio, digamos de 90 minutos, esto es 5,400 segundos, significaría un consumo de 108,000 megabytes, o 108 gigabytes, o 108,000 millones de caracteres. Al DVD más amplio en la actualidad le caben 10 gigabytes por los dos lados. Es obvia la razón por la que se requieren técnicas de compresión. La representación digital de un minuto de música en formato no comprimido de cualquier disco compacto musical —44,100 muestras por segundo, 16 bits por muestra— ocupa más de 9 megabytes, o 9 millones de caracteres. Tratar de descargar música o video de una biblioteca con estos volúmenes sería absolutamente impráctico por el tiempo que el proceso consumiría.

La compresión de datos no es nueva; como ejemplo de ellas tenemos el código desarrollado por Samuel Morse para la telegrafía basado en “puntos y rayas” eléctricas. Otro ejemplo interesante lo constituye el sistema Braille para lectura de invidentes: con arreglos de 2 x 3 puntos

se puede representar y leer cualquier texto. Como algunas de las combinaciones representan palabras y no sólo caracteres, se logra una reducción adicional del 20% en espacio.

Si bien el bibliotecario no tiene por qué conocer al detalle esas técnicas es importante que sepa de su existencia, de su importancia y, sobre todo, de sus consecuencias, y esté al tanto de sus avances y posibilidades. Ello le permitirá ir evolucionando con estos desarrollos y tomar mejores decisiones a la hora de decidir cuál formato usar para construir una colección dada. Recordemos que los formatos mencionados no son la última palabra; van evolucionando rápidamente y es necesario entonces moverse junto con ellos, ya que ninguno ha prevalecido a lo largo de muchos años.

A guisa de resumen es conveniente examinar, aunque sea someramente, cuáles son las técnicas de compresión más estudiadas y desarrolladas en la actualidad, y que fundamentan algunos de los formatos más utilizados hoy en día y serán la base de los formatos del futuro cercano. Cabe mencionar que prácticamente todas estas técnicas se basan en principios y algoritmos matemáticos que se le aplican a un proceso real.

Toda la técnica de compresión en general se basa en el principio de que: dado un algoritmo o proceso de compresión al cual entra un documento “X” obtendremos un documento comprimido “Y” que ocupa menos espacio que el original; si aplicamos el proceso inverso, al pasar por la técnica el documento comprimido “Y” deberemos obtener el documento original “X” *idéntico*. Si esto se cumple, se denomina a ésta técnica “sin pérdida” —non-lossy—.

En un segundo tipo de técnica, el archivo resultado, ya comprimido, ha perdido algo de su calidad original, y al aplicar un proceso inverso NO se tendrá el archivo original del cual se partió. *A cambio de ello se logran mayores niveles de compresión*. A este segundo tipo de técnica se le denomina “con pérdida” —lossy—. El archivo ya comprimido no contiene la misma información que el original, y en un proceso inverso no se podrá tener el archivo idéntico del cual se partió, sino sólo una aproximación.

Ninguna técnica es mejor que otra por sí misma; depende de la aplicación. Por ejemplo, en compresión de voz en un teléfono, hay

frecuencias que no son audibles para el oído humano, así que el eliminar esas frecuencias al momento de comprimir no tiene mayor efecto sobre la reconstrucción de un original no idéntico para la persona que escucha y por lo tanto no importa que la reconstrucción no sea idéntica al original. En el caso de un texto, es indispensable a la hora de reconstruir el archivo que el texto descomprimido sea idéntico al original. Por lo mismo, en algunos casos se evalúa un algoritmo de alguna manera y en otras de forma distinta, dependiendo del contexto y uso del documento a comprimir.

Estas técnicas o algoritmos, a cambio de la pérdida, o no, de fidelidad con el original, entregarán relaciones de compresión de mayor o menor grado; es decir, factores de compresión de dos a uno, de cuatro a uno, dieciséis a uno, etcétera, lo cual se convierte en un elemento de medición y calificación del algoritmo. Otras variables que sirven para calificar una técnica dada son la cantidad de memoria que consume en el computador a la hora de efectuar el proceso, la velocidad a la que se puede comprimir masivamente un conjunto de documentos, etcétera. Al existir entonces muchas técnicas propuestas para comprimir un tipo de documento en particular, es necesario estar consciente de las ventajas y desventajas de alguna técnica empleada en función del documento a comprimir. Entre las técnicas más utilizadas y estudiadas hoy en día distinguimos: [Sayood, 2000]

- Compresión fractal.
- Codificación de Shannon-Fano y Huffman.
- Técnicas de diccionario.
- Compresión LZ.
- Codificación aritmética.
- Código predictivo.
- Cuantización escalar y vectorial.
- Algoritmos de compresión tipo *jpeg*.
- Grupo 4

Ésta no es, ni con mucho, una lista exhaustiva; simplemente la quiero utilizar a guisa de ejemplo para ilustrar algunos de sus rasgos más representativos, sus ventajas y desventajas, y para conocer un poco acerca

de las técnicas más utilizadas en relación con los documentos en bibliotecas digitales, sean éstos textos, imágenes, audios o videos.

La compresión fractal se basa en técnicas para ir retirando elementos redundantes de la representación de una imagen. En ella, en vez de ir almacenando para cada pixel (punto) su color y posición, se crea un conjunto de instrucciones que permita reconstruir toda una parte parecida de una imagen. La teoría de los fractales no es reciente; sin embargo su uso en técnicas para trabajo con imágenes se remonta a 1977 con la obra de Benoit Mandelbrot *La Geometría Fractal de la Naturaleza*. En ella se sientan las bases para el manejo de imágenes con base en un conjunto de instrucciones. Hoy en día, muchos de los juegos electrónicos para computadora generan sus escenografías y personajes virtuales basados en técnicas de fractales. De manera inversa, una imagen “real” puede ser reducida a ciertos grupos de instrucciones, aunque sea parcialmente, y no a descripciones punto por punto. Esta técnica, hasta cierto nivel de compresión, puede resultar “sin pérdida”; más allá de ese nivel se logra mayor compresión, pero comienza a tener pérdida. [May, 1996]

A partir de sus obras, Claude Shannon³¹ es considerado como el primero que unió los trabajos anteriores en algo coherente que fue denominado la “teoría de la información”, de la cual parten la mayoría de las técnicas actuales. La técnica de compresión conocida como Shannon-Fano, fue desarrollada por Shannon en los laboratorios Bell y por Ugo Fano en el Instituto Tecnológico de Massachusetts (MIT) y presentada en 1948.³² Ésta se basa en una técnica de medición y compresión de datos en forma estadística que brinda una reducción considerable en datos de tipo alfabético. Esta técnica fue perfeccionada por David Huffman en los años cincuenta y es ampliamente utilizada hoy en día como algoritmo de compresión de textos, así como base para comprimir otro tipo de datos, como modems, faxes, redes de cómputo, etcétera.

31 “Bevis King’s Video Formats”. Disponible noviembre 3, 2005 en: <http://www.ee.surrey.ac.uk/Contrib/WorldTV/video.html>

32 Shannon, Claude. 1948. “A Mathematical Theory of Communication”. En: *The Bell System Technical Journal*. Bell Laboratories, (1948). pp. 379-636.

En esencia, esta técnica se basa en la probabilidad de frecuencia de ocurrencia de un carácter dentro de un texto en una lengua dada. Esto se ha logrado contando la incidencia de letras en múltiples textos de un idioma, y asignándole códigos binarios más reducidos en número de bits que los usualmente ocho arrastrados en el código ASCII universal. De este modo, se reduce sensiblemente el número de bits utilizados para almacenar y reproducir un texto pictográfico en forma digital.

La técnica de compresión “LZW” —Lempel-Ziv-Welch— fue ideada por Jacob Ziv y Abraham Lempel, y publicada en los anales del Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) en 1977;³³ los refinamientos de Terry Welch al algoritmo para ser usado en imágenes fueron publicados en 1984 y de ahí tomó el nombre completo. Dependiendo de la época, se conocieron como técnicas LZ77, LZ78, LZ2, etcétera. El algoritmo LSW sustituye cadenas identificadas de caracteres —llamadas “cuerdas”— con códigos más simples; no se hace ningún análisis semántico del texto entrante, solamente se identifican nuevas cuerdas de caracteres y se sustituyen por nuevos códigos. Al ser éstos más simples y menores en tamaño que las cuerdas a las que sustituyen, ocurre la compresión. Esta técnica fue ampliamente utilizada y dio origen, entre otros, al formato “GIF” para imágenes, en el cual ciertas partes fueron patentadas por la empresa Sperry, quien a la larga se fusionó con Burroughs formando Unisys, hoy en día el titular de estas patentes.

La compresión aritmética es una versión más avanzada de las técnicas de Huffman. La diferencia básica consiste en que en la técnica de este científico se requieren potencias de números enteros de bits para poder hacer la representación de un carácter. En la compresión aritmética se usan potencias no enteras, es decir, números reales, lo cual hace todavía más eficiente el algoritmo de compresión, ya que se usan menos caracteres. La desventaja es la complejidad. Mientras que el algoritmo de Huffman es bastante simple y por lo mismo fácilmente programable, la compresión aritmética es mucho más compleja en su

33 Wyner, Aaron D. *The Significance of Shannon's Work*. Disponible noviembre 3, 2005 en: <http://cm.bell-labs.com/cm/ms/what/shannonday/work.html>

desarrollo y programación. Las técnicas de compresión de texto enunciadas anteriormente no tienen pérdida.



FIGURA 2.3 TRES IMÁGENES SUCESIVAS AL APLICAR UN ALGORITMO DE COMPRESIÓN-DESCOMPRESIÓN CON PÉRDIDA³⁴.

La técnica del “código predictivo” se basa en ocurrencias pasadas de algún tipo de datos para estimar la probabilidad de que ese dato se presente de nuevo; es decir, se comprime más lo que más se ha presentado. Ello permite una compresión más eficiente. Entre estas técnicas se distinguen la de “predicción con coincidencia parcial” —Prediction with Partial Match o PPM—, las transformadas de Burrows-Wheeler (BWT), la técnica CALIC —*Context Adaptive Lossless Image Compression*— o “compresión de imágenes sin pérdida por contexto adaptativo”.

La técnica de “cuantización escalar” se basa en el proceso de representación de un conjunto de valores de palabras código que a su vez representa un conjunto aún mayor de valores de palabras fuente. Ésta es una técnica muy eficiente, pero del tipo “con pérdida”. La “cuantización vectorial” es semejante pero trabaja con “bloques” de datos a la vez, llamados “vectores”.

³⁴ Ziv, Jacob, y Lempel, Abraham. 1977. “A Universal Algorithm for Sequential Data Compression”. En: *IEEE Transactions on Information Theory*, IT-23:3. (Mayo 1977).

La técnica de compresión tipo “JPEG” fue desarrollada por el “*Joint Photographic Experts Group*”, un comité que definió el estándar. Se diseñó para la compresión de imágenes fijas: fotografías, mapas, carteles, etcétera, ya sea a color o en blanco y negro. Es una técnica “con pérdida” pero que logra sensibles factores de compresión de una imagen hasta en un séptimo del tamaño del original en bytes; es decir en volumen, con sólo una pérdida de 25% en calidad. Fue diseñada para distribuir imágenes vía la Red, donde la pérdida para ese propósito y por ese medio era bastante aceptable. Compárense los tamaños de una misma imagen en un formato, digamos, “TIFF”, el cual es “sin pérdida”, o “GIF”, con mínima pérdida, y podrá observarse la enorme diferencia entre el volumen de unos y otro, mientras que la calidad no disminuye en la misma proporción. Pero no recomendaríamos “JPEG” con fines de preservación de imágenes.

Otro estándar de compresión internacional adoptado por muchas bibliotecas digitales tiene que ver con el envío de datos a través de la red. En la actualidad el más utilizado de éstos es el conocido como “Grupo 4” —Group 4 for 1-bit images—. Este estándar fue introducido originalmente en 1988 para módems y faxes por el CCIT, —*Comité Consultatif International Télégraphique et Téléphonique* o *International Consultative Committee on Telecommunications and Telegraphy*—. El CCITT cambió su nombre en 1993 a ITU-T —*International Telecommunications Union - Telecommunications Standardization Sector*—. Las especificaciones se denominan “Facsimile Coding Schemes and Coding Control Functions for Group 4 Facsimile Apparatus”, ITU-T T.6.

Como puede verse, la compresión y tipo de técnica a utilizarse ha de ser seleccionada de acuerdo con un criterio que logre el balance entre el propósito que buscamos y su eficiencia al almacenarse y distribuirse. No existe la técnica perfecta de compresión que maximice todas las variables y por ello debe ser analizada y seleccionada de antemano de acuerdo con un propósito definido y un criterio. El bibliotecario debe estar al tanto siempre de cuál es el formato en que cada una de sus colecciones es digitalizada y de cuáles son sus características de compresión y uso de almacenamiento. Con esto hemos concluido la revisión de los puntos importantes que deben ser considerados

por una biblioteca al momento de convertir materiales, ya existentes en formatos tradicionales, hacia formatos digitales.

Como hemos podido observar a lo largo de este capítulo, los conceptos relacionados con la digitalización son numerosos y variados, pero una vez que sus parámetros se comprenden esto ayuda enormemente al bibliotecario profesional a tomar decisiones tendientes a crear colecciones digitalizadas de una forma adecuada, y por lo tanto es indispensable establecer previamente a estos proyectos la serie pertinente de criterios para realizarlos. Como un gran resumen de este capítulo pueden verse en el anexo 4 unas tablas que compendian las características más relevantes de la digitalización, así como algunos vínculos muy interesantes al respecto.

Capítulo 3

Colecciones de documentos originalmente digitales

"El mundo produce entre uno y dos exabytes de información original por año. Un exabyte es un trillón, o 10^{18} , o 1,000,000,000,000,000,000 de bytes".

Peter Lyman y Hal Varian.³⁵

3.1 PUBLICACIÓN ELECTRÓNICA O DIGITAL

Carol Tenopir,³⁶ una de las notables estudiosas de las publicaciones electrónicas, ha señalado que:

en la vertiginosa carrera hacia un mundo de información digital rara vez nos detenemos a considerar el efecto a largo plazo para las bibliotecas, los académicos y los estudiantes. Más raro todavía es reflexionar acerca del efecto a largo plazo que los cambios en los medios de comunicación académica traerán en el aprendizaje y en la comprensión de contenidos.

Marshall McLuhan³⁷ ya había presentido en 1964 los impactos entre las variaciones de los medios y los significados y su impacto en la sociedad a través de su famoso postulado *el medio es el mensaje*.

Si bien los trabajos de McLuhan se derivaron de los medios masivos de comunicación: radio, TV, prensa, —*no había Internet entonces*— y no con publicaciones académicas, es válido plantearse la pregunta con el advenimiento de la Red. Las variaciones en los medios de comunicación de los materiales académicos: revistas, reportes, bancos de datos, etcétera, ¿implican cambio en cómo son usados o entendidos, o el

35 Lyman, Peter y Varian, Hal. 2000. "How Much Information?". En: *Journal of Electronic Publishing*. Diciembre, 2000, vol. 6:2. ISSN 1080-2711. Disponible noviembre 3, 2005 en: <http://www.press.umich.edu/jep/06-02/lyman.html> Véase También Anexo 1

36 Tenopir, Carol. 2003. *Electronic Publishing: Research Issues for Academic Librarians and Users*. Disponible noviembre 3, 2005 en: http://www.findarticles.com/p/articles/mi_m1387/is_4_51/ai_105046543/pg_1

37 McLuhan, Marshall. 1964. *Understanding Media: The Extensions of Man*. New York : McGraw-Hill.

contenido permanece separado del medio? Dicho de otra forma ¿hay diferencias sustanciales para el usuario entre los materiales académicos entregados en papel y los de forma electrónica?

La misma Carol Tenopir agrega las siguientes preguntas:

- ¿Cómo afecta el medio de distribución las preferencias, comprensión o uso de los contenidos?
- ¿Las personas buscan, comprenden y usan la información entregada en medios digitales de manera diferente a la entregada en papel?
- ¿Las fuentes en estos nuevos medios de distribución producen alguna diferencia en la manera en que las personas crean conocimiento a partir de esas fuentes?
- ¿La elección del formato de los productos de información por parte de la biblioteca, afecta a los usuarios?

Esta autora concluye que éstas son algunas de las preguntas clave que los bibliotecarios deberán responder en la siguiente década en torno a las publicaciones electrónicas. Ella misma agrega que para poder responderlas será necesario plantear y responder muchas más preguntas previas. En este sentido desarrollo el análisis de este capítulo, pero planteo a la vez algunas de las interrogantes y de las reflexiones que me van surgiendo al realizar el análisis.

En primer término, como en secciones anteriores, conviene definir qué se considera como una publicación “electrónica” o “digital”. Hablando con propiedad el término más correcto debiese ser *publicación digital*, como algunos autores la han denominado, ya que la estructura fundamental de la información se encuentra en formatos digitales y se distribuye y obtiene por medios o soportes electrónicos u ópticos. No obstante, el término *publicación electrónica*, que ya rebasa las tres décadas de acuñado, sigue siendo ampliamente utilizado en la literatura actual y se rehúsa a cederle su lugar al mencionado previamente; por lo mismo, me seguiré refiriendo a este tipo de publicaciones como *publicaciones electrónicas*. Históricamente éste es un término que se ha venido conformando a lo largo de las últimas tres décadas del siglo xx y que todavía se encuentra en plena evolución. Como muchos de

los elementos que hemos estado revisando, en distintos momentos, el término “publicación electrónica” ha significado algo distinto para diferentes grupos de personas. Observado por épocas, a principios de los años setenta, *publicación electrónica* para las bibliotecas eran los catálogos y los índices y resúmenes en publicaciones periódicas secundarias distribuidas en cintas magnéticas o accesibles en línea a través de grandes *mainframes* o computadoras centrales. En los ochentas, publicación electrónica se definía como una publicación —libro, revista, ficha catalográfica, bibliografía, etcétera— que era impresa con ayuda de una computadora. A mediados de esa década la empresa “*Learned Information*” lanzó “*The Electronic Magazine*”, una base de datos totalmente electrónica. Noticias de Europa y los Estados Unidos eran reunidas en microcomputadoras y teletransferidas a Oxford, donde eran tratadas y combinadas con otras noticias y agregados recibidos de otros lugares; ahí la revista se iba formando con noticias, artículos cortos, revisiones de libros, publicidad, etcétera. Después de su proceso estos nuevos elementos eran cargados en discos en Roma, y se ofrecía la revista para su consulta en línea.

Para los noventa, se observan dos variantes; por un lado, el advenimiento del CD-ROM de forma masiva y económica les permitió a los editores de publicaciones periódicas secundarias, bancos de datos, catálogos, etcétera, distribuirlos ampliamente con este medio para ser usados en bibliotecas por medio de sus computadoras personales o “torres” de discos compactos. Poco después veríamos también aparecer los CD con textos completos e información primaria. La segunda variante de estas publicaciones se dio también a principios de la década con la llegada de Internet, a través de una de sus primeras y simples funciones básicas, el “FTP”, —*File Transfer Protocol*—. Éste era un protocolo normalizado que simplemente permitía el intercambio de archivos de computadora entre equipos de distinta naturaleza. Ello hacía posible que los científicos intercambiaran sus archivos con textos provenientes de su procesador de palabra o sus archivos de datos. Debe recordarse que no existía la “*World Wide Web*” ni el lenguaje HTML para publicación de textos. Muchos documentos fueron “publicados” así, como el simple intercambio de archivos electrónicos. Véase la gráfica en la figura 3.1: *distribución de e-journals por Internet* y podrá

observarse esta evolución. Antes de 1995, prácticamente todas las “publicaciones electrónicas” de este tipo se distribuían vía correo electrónico o *Gopher*, para aprovechar el protocolo *FTP*.

Ya para mediados de la década de los noventa, con el crecimiento de la red mundial *www*, comenzó a darse por sentado para un gran sector de los usuarios que “publicación electrónica” representaba a las revistas y otros materiales semejantes que se publicaban vía la red, y empezaron a ser usados como sinónimos y a convertirse en el arquetipo de estas publicaciones. En mucha de la literatura del último lustro del siglo XX, podemos observar este fenómeno, aunque en realidad, debemos estar conscientes de que, objetivamente, abarca más cosas. Boyce y Dalterio definieron así en 1996 la problemática de los diversos elementos involucrados en la publicación electrónica: [Boyce y Dalterio, 1996, p. 42]

Mucha de la tensión entre autores y editores acerca de la publicación electrónica proviene de una comprensión fragmentada de lo que toma tener un producto completo. No hay ni siquiera un consenso general en lo que ‘*publicación electrónica*’ significa. Los diversos grupos involucrados tienden a enfocarse sólo en los aspectos que le son familiares. Los autores se concentran en la preparación electrónica de manuscritos y su sometimiento a revisión. Los lectores se enfocan en la recuperación de información vía Internet. Los bibliotecarios se abocan en la entrega de información al usuario, pero a menudo ignoran la colección o archivo electrónico. Los editores se preocupan del manejo de los manuscritos electrónicos; registro, formateo, tipografía y producción de versiones adecuadas para entrega al usuario final. También se preocupan de recolectar las utilidades necesarias para mantener su operación financieramente viable.

Ellos mismos definieron también los pasos necesarios para producir una revista electrónica: Nótese que la definición sigue siendo de 1996: [Boyce y Dalterio, 1996, p. 43].

- Preparar y someter el material manuscrito. Después de escribir el documento, los autores están por lo general impacientes por distribuir sus resultados a otros colegas. La espera por arbitraje y registro puede parecer interminable.
- Arbitraje por pares para asegurar calidad científica. Si bien algunos tienen discusiones sobre este punto, hay una clara distinción entre

artículos de conferencias sin arbitraje y los de revistas científicas que sí han pasado por este proceso. Desde muy al principio, ICSU —International Council of Scientific Unions— y Unesco recomendaron muy puntualmente que el arbitraje no debía ser menor en las revistas electrónicas que en las de papel. [Joint ICSU Press / UNESCO, 1996].

- Registro, formado y tipografiado. Muchos autores no aprecian adecuadamente la importancia de estos pasos para mejorar la precisión y utilidad de la información transferida. La claridad en la escritura y buena legibilidad de las páginas son características importantes para una buena revista.
- Preparación de la base de datos. La base de datos es un elemento crucial del sistema de diseminación electrónica de información. Su preparación incluye el contar con buenas herramientas de almacenamiento y recuperación.
- Producción y diseminación en múltiples formatos. El paso final en el sistema de entrega consiste en proveer la información en un formato adecuado al medio de entrega y a las necesidades del usuario.
- Archivo. El almacenamiento permanente de artículos publicados para su uso futuro es una tarea continua. La evolución de las tecnologías de almacenamiento y la eventual transferencia de materiales archivados a nuevas generaciones de medios de almacenamiento vendrá a ser una tarea creciente para editores y bibliotecarios.

Si se analiza con cuidado esta última definición de *publicación electrónica* puede observarse la falta de precisión que ella encerraba. Se da por hecho que *publicación electrónica* y *revista electrónica* son sinónimos. Pero si se observan los “pasos” para el desarrollo de la publicación, pueden advertirse los estilos propios de la época, muy al estilo de publicaciones tradicionales simplemente convertidas a un formato electrónico. Falta toda una serie de elementos de marcado, estandarización, hipervínculos, etcétera, que ya pueden observarse fácilmente hoy en día.

Una traducción de una definición simple y adecuada de lo que es una publicación electrónica en la actualidad puede obtenerse del diccionario en red *Computeruser*:

Publicación Electrónica: La producción de documentos para ser vistos en una pantalla de computadora u otro dispositivo semejante, los cuales pueden no haber sido editados originalmente en papel. Los documentos así

publicados pueden estar en CD-Rom, disco flexible, o encontrarse en redes de computadoras como el Internet, y además de texto e ilustraciones, pueden contener trozos de video o sonido, gráficas animadas e hipervínculos.

Nuevamente hay que destacar aquí que el término “documento” debe entenderse en esta definición en su acepción más amplia: libro, revista, manuscrito, mapa, modelo, audio, fórmula, plano, video, etcétera.

El Diccionario de Bibliotecología y Ciencias de la Información ODLIS define publicación electrónica como:

la publicación en formato digital de libros, revistas, bases de datos bibliográficas, y otros recursos de información, por lo general sobre CD-Rom o en-línea vía la Internet, para usuarios locales, suscriptores, o clientes de menudeo, con o sin una contraparte impresa. Sinónimo de ‘publicación asistida por computadora’”. [ODLIS, 2002]

Al igual que la evolución de bibliotecas electrónicas a bibliotecas digitales, algunos autores están usando actualmente el término “publicación digital” para proyectar la sensación, real o no, de que están usando conceptos más innovadores en publicaciones de este tipo.

3.2 REVISTAS CIENTÍFICAS

Al analizar el estado del arte de las publicaciones electrónicas en la actualidad, conviene empezar por las publicaciones periódicas. Si bien la publicación por antonomasia es el libro, y casi todos los estudios bibliográficos comienzan con ellos dada su importancia, en el caso de las publicaciones electrónicas la revista tiene una mayor historia y desarrollo que el libro, y por lo tanto es el primer tipo de publicación que conviene estudiar ya que proporciona más elementos de análisis y una mayor ventana de tiempo que nos permite tratar de establecer tendencias.

En 1980, la Universidad de Birmingham y la Universidad Tecnológica de Loughborough en Inglaterra anunciaron conjuntamente el pro-

yecto “BLEND” —*Birmingham and Loughborough Electronic Network Development*—. Los objetivos del proyecto eran estudiar los problemas que surgen al establecer una comunidad de información y una revista electrónica. El concepto de “revista electrónica” fue descrito como:

la utilización de una computadora para ayudar los procedimientos normales por los cual un artículo es escrito, es arbitrado, aceptado, y publicado. El autor, árbitros, editor y alternativamente los lectores pueden tener así acceso al texto de los artículos desde sus computadoras.

En 1987, La Universidad de Siracuse, N.Y. publicó a través de un proyecto especial la revista *New Horizons in Adult Education*, que fue la primera revista arbitrada distribuida vía la *Internet*, estaba en texto plano *ASCII*, era gratuita y se distribuía vía el servidor de *BITNET*. En 1990 la siguió *Postmodern Culture*, otra revista arbitrada, también en *ASCII* a la usanza de la época y distribuida por la red; todavía se produce hoy en día y está disponible en el sitio *Web* del proyecto *Musa*: <http://www.iath.virginia.edu/pmc/contents.all.html>

En 1992 apareció la que sería la primera revista electrónica arbitrada, en texto completo, y que incluía ya gráficas: el *Online Journal of Current Clinical Trials*, un esfuerzo de *OCLC* —*On line Computer Library Center*— muy al principio para apoyar el desarrollo de una revista por suscripción, electrónica que no estuviera también en papel. Fue un proyecto muy avanzado para su tiempo, ya que no existía la *Web*, ni los *buscadores* y requería por tanto de una interfaz especial para poder visualizarse. No obstante, sirvió como base para el proyecto *ECO* —*Electronic Collections Online*— de esta empresa.

En 1993 surgió *JSTOR* —*Journal Storage Project*— el cual se ha ido consolidando como el mayor proyecto de digitalización de archivos retrospectivos de revistas impresas que va hacia una versión electrónica. A la fecha, numerosos títulos han sido digitalizados e incorporados al proyecto. Puede abundarse en la página oficial del sitio: <http://www.jstor.org/>

Cabe resaltar aquí que el concepto de *revista científica* debe tomarse en su acepción más universal de la ciencia, y no tan sólo pensarse

en las ciencias “duras”, sino también las que abarcan las ciencias sociales, las humanidades y las artes. Es decir, el concepto se aproxima más a lo que sería una *revista académica*, aunque tiende a describirse en la literatura como revista científica.

Con el advenimiento de la *Web*, prácticamente cualquiera ha podido convertirse en editor y ofrecer “revistas” en la red mundial, que poco o nada tienen que ver con las ediciones cuidadas, frecuentes, arbitradas, originales, etcétera, que caracterizan a una verdadera publicación.

Desde la década de los ochenta se ha venido hablando en la literatura especializada de la denominada “crisis de las revistas académicas”, la cual ha sido motivada principalmente por dos problemas interrelacionados: por un lado los cada vez más elevados precios de las revistas ante un mercado reducido y sin competencia, y por el otro, la gran demora entre la fecha de entrega del artículo y su fecha de publicación. Si bien no son los únicos problemas implicados, sí se considera que son los dos que más han incidido en el medio.

Con el advenimiento de las revistas digitales prototipo en la red, de las cuales ya hemos mencionado algunas, se ha venido modificando cada vez más el concepto de distribución y acceso a la información. Entre las ventajas más atractivas de este nuevo medio se distinguen las siguientes:

- Tiempo mucho menor de edición entre la entrega del artículo y su distribución.
- Considerable reducción de costos de edición, ya que los elementos de papel y correo no existen.
- Otras capacidades adicionales del medio digital.

Continuemos con la definición formal de la revista electrónica o digital. Según el *Glosario de bibliotecología y Ciencias de la Información de la American Library Association*³⁸ Una publicación periódica es:

38 American Library Association (ALA). 1983. *The ALA Glossary of Library and Information Science*. American Library Association, Chicago, IL.

Una serie que aparece o pretende aparecer indefinidamente en intervalos regulares, por lo general con más frecuencia que anualmente, y en donde cada fascículo es numerado o fechado consecutivamente; normalmente contiene artículos independientes, historias u otros escritos. No se consideran en esta definición los periódicos que diseminan noticias en general, así como las memorias, anales, u otras publicaciones de organismos corporativos primariamente relacionadas con sus reuniones.

De acuerdo con lo anterior, si unimos las dos definiciones previas y se le agregan algunas precisiones finas, podemos aventurarnos a hacer una definición de revista científica electrónica más o menos actualizada:

Una serie en la que el contenido es sujeto de arbitraje editorial por expertos, que aparece o pretende aparecer indefinidamente en intervalos regulares, por lo general con más frecuencia que anualmente, y en donde cada fascículo es numerado o fechado consecutivamente, aunque puede tener actualizaciones y comentarios; normalmente contiene artículos independientes, historias u otros escritos que han sido preparados para ser vistos en una pantalla, y pueden o no tener equivalentes impresos en papel. Pueden estar publicadas en medios electrónicos portátiles, como discos magnéticos u ópticos y memorias, o bien encontrarse en redes de computadores como la Internet, y además de texto e ilustraciones, pueden contener otras modalidades documentales, como audio o video, gráficas animadas, hipervínculos, simulaciones, etcétera.

Para la primera etapa de las revistas electrónicas es muy útil revisar la evolución del *Directorio de Revistas Electrónicas, Boletines y Listas de Discusión Académica —Directory of Electronic Journals, Newsletters and Academic Discussion Lists—*, publicado por la *Asociación de Bibliotecas de Investigación —Association of Research Libraries, o ARL—*, el cual fue editado entre 1991 y 1997. [Mogge, 1999]. Sigue disponible en la actualidad dentro del sitio “Web” de la ARL [Association of Research..., 2000].

En 1990, la ARL y la Universidad Estatal de Carolina del Norte convocaron a una reunión entre los principales editores de revistas académicas, así como entre los bibliotecarios asociados, con objeto de analizar esa emergente tecnología. Ocho revistas estuvieron representadas, de las cuales, cinco continúan hasta nuestros días y son ya “incunables”

de las revistas electrónicas. Se observa que esas revistas iniciales eran de origen universitario, aunque sin patrocinio de su institución. Curiosamente, las áreas de esas revistas eran principalmente las humanidades y las ciencias sociales, ya que en ese entonces el texto plano era el método predominante de envío de información en la red, y este tipo de material se prestaba a ello. Como ya hemos mencionado, la inmensa mayoría de las revistas se entregaba vía FTP o por correo electrónico asociado a listas de entrega. La *“World Wide Web”* no sería una realidad práctica sino hasta 1994, cuando apareció el navegador *“Mosaic”* y poco después *“Netscape”*, los cuales permitieron el uso de texto, imágenes e hipervínculos simultáneamente.

La primera edición del directorio, en 1991, consignaba 110 revistas y boletines electrónicos, así como 517 listas académicas de discusión. Para la tercera edición en 1993, las revistas eran 240 y las listas 1,152. Para 1994, se observan ya las primeras 11 revistas diseñadas específicamente para tomar ventaja de las características de la *Web*, y se distribuían ya por este medio 60, de un total de 443 revistas y boletines consignados. El número de listas, y memorias de conferencias electrónicas ascendió a 1,758. En 1995, con la quinta edición del Directorio, 675 revistas y boletines eran consignados. Aunque el número de revistas verdaderamente académicas era de 142, en el resto se notaba ya la potencialidad multimedial de la *Web*, y múltiples *“magazines”* comenzaron a aparecer, dadas las capacidades de comercialización y divulgación de la *“Web”*. Las memorias de conferencias electrónicas llegaron a 2,480, diez veces más que la primera edición. Para la sexta edición en 1996, de un total de 1,689 revistas y boletines electrónicos consignados, más del 90 % eran ya distribuidos vía la *Web*, y sólo el 14 % mantenía el correo electrónico como medio de distribución y el 17 % ofrecía acceso por FTP. Las memorias y listas llegaron a 3,118. En Diciembre de 1997, cuando apareció la última edición del Directorio, el monto era de 3,400 revistas y boletines, así como 3,800 memorias y listas.

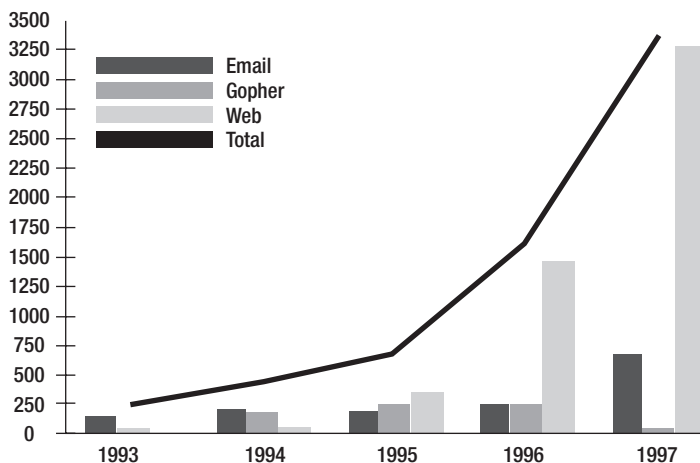


FIGURA 3.1: "DISTRIBUCIÓN DE E-JOURNALS POR INTERNET". TOMADA DE: ARL DIRECTORY OF SCHOLARLY ELECTRONIC JOURNALS AND ACADEMIC DISCUSION LISTS. 2000.

Considerando que ninguno de los títulos consignados en la tercera edición tenía una dirección URL, y sólo dos años después, en la sexta, más del 90 % de los títulos lo tienen, cabe notar el impacto y crecimiento que en ese periodo tuvo la *Web* como promotora de la publicación electrónica, en especial las revistas científicas electrónicas. He aquí una tabla con la serie que actualiza y complementa la gráfica anterior de revistas electrónicas:

MEDIO DE LA REVISTA	2001	2002	2003
Disco magnético	679	686	701
CD-ROM	2,898	3,292	4,160
en-línea	13,278	17,394	22,916

FIGURA 3.2 – NÚMERO DE REVISTAS ELECTRÓNICAS CON ISSN
TOMADA DEL: SITIO OFICIAL DE ESTADÍSTICAS DE LA ORGANIZACIÓN DEL ISSN.

Cabe resaltar que con cada edición el concepto de lo que debía incluirse en el *Directorio* se hacía más estricto. Para 1995 ya no fueron incluidos periódicos, gacetas, boletines de divulgación, publicaciones

periódicas secundarias, anuarios, etcétera. Si bien no todos los títulos incluidos en el *Directorio* fueron estrictamente académicos en el sentido recto de la palabra, es necesario resaltar que hubo un esfuerzo permanente de no incluir “literatura gris electrónica” en esa publicación, por lo que pueden inferirse interesantes estudios de la misma. Por ejemplo, un 30 % de las revistas en el *Directorio* llegaron a tener arbitraje, y sólo una cuarta parte de ellas se distribuía a cambio de una tarifa de suscripción. Como puede concluirse, no habían llegado las grandes firmas editoras de revistas; lo hicieron hasta 1996.

En efecto, en ese año aparecen las primeras revistas científicas electrónicas provenientes de las grandes casas editoriales internacionales. Así, las revistas de *Elsevier Science*, *Academic Press*, *Springer-Verlag*, *Chapman-Hall*, etcétera, comenzaron a aparecer en el escenario. En esencia, han sido hasta ahora versiones digitalizadas casi idénticas a las ediciones en papel, con algunos valores agregados.

TÍTULOS CON ARBITRAJE Y TOTALES

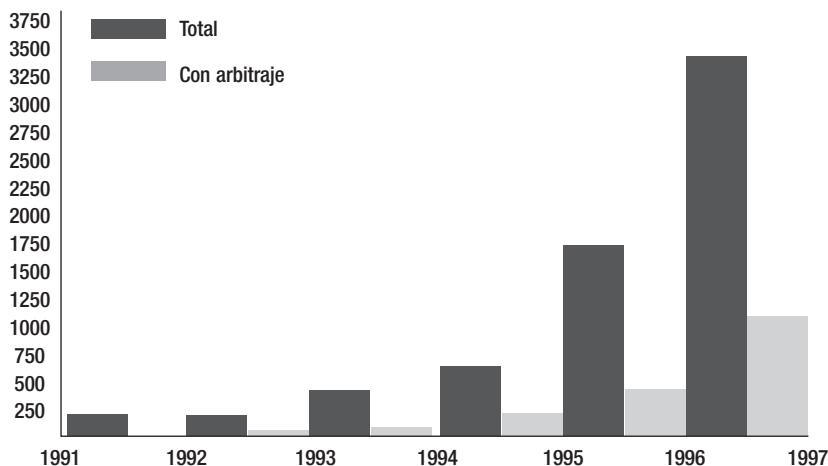


FIGURA 3.3: TÍTULOS DE REVISTAS ELECTRÓNICAS POR ARBITRAJE. TOMADA DE: ARL DIRECTORY OF SCHOLARLY ELECTRONIC JOURNALS AND ACADEMIC DISCUSSION LISTS. 2000.

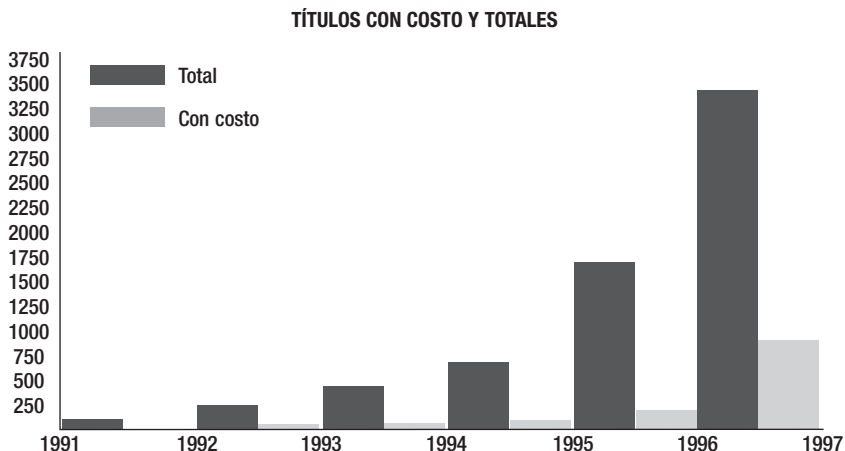


FIGURA 3.4: TÍTULOS DE REVISTAS ELECTRÓNICAS POR TIPO DE COBRO. TOMADA DE: ARL DIRECTORY OF SCHOLARLY ELECTRONIC JOURNALS AND ACADEMIC DISCUSSION LISTS. 2000.

De acuerdo con Kling y McKim, se distinguen cuatro tipos principales de revistas electrónicas: [Kling y McKim, 1997]

e-journals “puros” — aquellos cuya publicación solo se realiza en medio digital, sin mediar publicación previa en papel, como la *Revista Digital Universitaria* —RDU— de la UNAM o el *Journal of Electronic Publishing* —JEP—.

e-p-journals — aquellos que se distribuyen primordialmente en forma electrónica, pero de los que puede haber limitadas copias en papel. Por ejemplo *The Journal of Artificial Intelligence Research* y *The Electronic Transactions on Artificial Intelligence*.

p-e-journals — aquellos que son distribuidos primordialmente en papel, pero que existen también en versión digital. Como ejemplos tenemos *Science*, *Investigación Bibliotecológica*, *Physical Review*.

p+e-journals — Aquellos que de inicio son lanzados con versiones en papel y electrónicas igualmente importantes. Como ejemplo está *The American Chemical Society's Organic Letters*.

Dados los avances tecnológicos, yo agregaría una quinta categoría, la denominada **blogs**. La palabra proviene de la contracción de un

término de la red: *Web Log —contenidos personales en la Red—*. En esencia, un “blog” es básicamente una revista electrónica o serie de artículos que se distribuyen por la red mundial *www*, con la característica distintiva de que su información se actualiza constantemente; es decir, no queda estática como en una revista “típica”. Algunas llegan a actualizarse diariamente, con nuevos comentarios y “vínculos”. Aunque no es una modalidad muy utilizada en revistas científicas propiamente dichas, ha ganado mucha popularidad en revistas de difusión y divulgación científica y tecnológica.

¿Cuáles son los valores agregados que las revistas electrónicas deben tener? Nuevamente, para contestar a esta pregunta, debe acudirse a los diversos sectores que tienen que ver con este tipo de publicación para definir la amplia variedad de puntos de vista que existen al respecto. Rusch-Feja y Siebeky realizaron un interesante estudio sobre la percepción de un grupo de investigadores en el Instituto *Max Planck* de Alemania acerca de las revistas electrónicas. Dicho estudio, para simplificarlo, fue acotado a las revistas de tres grandes firmas: *Elsevier*, *Academic-Press* y *Springer-Verlag*. La muestra fue de aproximadamente 11% de los académicos del Instituto y si bien no es muy actual ilustra perfectamente lo que los autores buscan al publicar en revistas académicas, electrónicas o no. [Rusch-Feja y Siebeky, 1999].

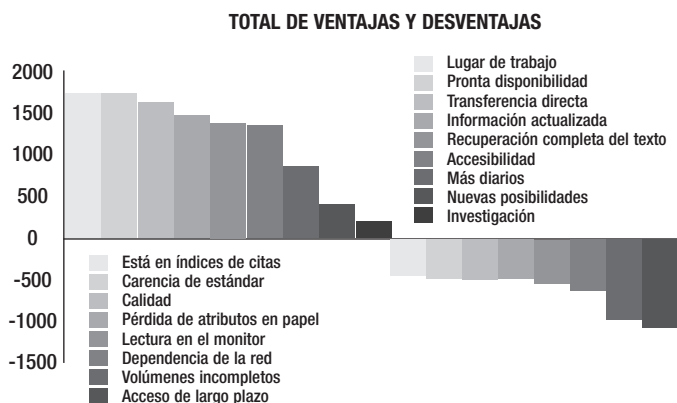


FIGURA 3.5: PERCEPCIÓN DE LOS USUARIOS ACERCA DE LAS VENTAJAS Y DESVENTAJAS DE LAS REVISTAS CIENTÍFICAS ELECTRÓNICAS. TOMADO DE: RUSCH-FEJA, D. Y SIEBEKY, U. 1999. “EVALUATION OF USAGE AND AND ACCEPTANCE OF ELECTRONIC JOURNALS. D-LIB MAGAZINE, VOL. 5, NO. 10”

Entre los resultados presentados es interesante analizar la gráfica de la figura 3.5 acerca del total de ventajas-desventajas percibidas por los encuestados. La pregunta se hizo pidiendo que se calificara cada aspecto desde *muy ventajoso* hasta *muy desventajoso* pasando por una gama de valores intermedios.

En el Reino Unido, el *Baseline Focus Group*, basado en un proyecto llamado *Superjournal*, realizó también un estudio en 1999 acerca de las ventajas o comodidades valoradas por los usuarios en las revistas electrónicas de varias Universidades inglesas: *Oxford, Cambridge, Leeds, Birmingham, Ulster, Warwick, Sussex, UCL —University College London—, NIMR —National Institute for Medical Research—, DMU —De Montfort University—, LSE —London School of Economics and Political Science—*. [Baseline Focus Group, 1999]. Los resultados de este muy interesante estudio reflejan los puntos de vista de una gama muy amplia de interesados. El estudio fue realizado por medio de encuestas a través de la *Web*, y puede observarse en una tabla resumen que presento a continuación:

Universidades	Cambridge	Oxford	UCL	NIMR	Leeds	Birm.	Ulster	Warwick	Sussex	DMU	LSE
Acceso											
Conveniencia, acceso en escritorio	X	X	X		X	X	X	X	X	X	X
Rápido y fácil acceso	X	X	X	X	X	X	X	X	X	X	
Acceso garantizado	X	X	X	X	X	X					
Acceso a toda hora			X					X	X		
Acceso desde cualquier parte	X		X			X	X		X	X	
Contenido											
Amplio rango de títulos	X	X	X	X	X			X		X	
Títulos extra de las suscripciones				X			X		X	X	
Años retrospectivos deseados	5 a 10		2 a 3	5+	5 a 10	30		15		5 a 10	
Búsqueda											
Búsqueda de gran rango de títulos		X	X	X	X	X	X	X	X	X	
Fácil y eficiente búsqueda	X	X	X	X	X	X	X	X	X	X	X

Bibliotecas y publicaciones digitales

Universidades	Cambridge	Oxford	UCL	NIMR	Leeds	Birm.	Ulster	Warwick	Sussex	DMU L	LCE
Búsqueda por autor y palabras clave	X	X	X	X	X	X				X	X
Búsqueda dentro del texto completo	X					X	X	X	X	X	X
Ligas											
Desde las referencias al artículo	X	X	X	X	X		X	X		X	X
Hacia material relacionado						X	X	X	X	X	
Hacia bancos de datos	X	X	X	X	X						
Hacia fuentes originales							X	X	X		
Hacia nuevas versiones del artículo	X		X		X						
Hacia el e-mail del autor o sitio Web		X	X	X	X						
Notas pie de página expandidas			X			X			X		
Integra acervo											
Todas las revistas en un sólo sitio	X	X	X	X	X	X		X	X	X	
búsqueda, abstract, artículo a la vez		X	X	X	X			X	X	X	X
Multimedios											
Datos extras para profundizar		X	X	X	X	X	X			X	X
Gráficas en tres dimensiones	X	X	X	X	X						
Video o película			X			X	X				
Audio o música		X				X		X			
Animaciones		X			X						
Oportunidad											
Revistas antes que en papel		X	X	X	X			X		X	
Otros											
Servicio de alerta			X	X	X	X		X		X	
Archivo electrónico de mi interés	X	X	X		X		X			X	X
Buena calidad de impresión/lectura	X	X	X		X		X		X		
Acomodo propio de títulos de interés	X	X	X			X					
Copiado de bibliografías propias		X								X	X
Crear/actualizar listas de estudiantes						X				X	X
Licencia de imágenes para docencia	X		X								
Foros de discusión										X	X

Las primeras cuatro, *Cambridge, Oxford, UCL y NIMR* forman un *Cluster* o agrupamiento de títulos especializados en Genética Molecular y Proteínas. *Leeds* un agrupamiento en Química de materiales, *Birmingham, Ulster, Warwick y Sussex* en Comunicaciones y estudios de la cultura; *DMU y LCE* en Ciencias políticas.

Varios análisis interesantes pueden desprenderse del estudio de esta tabla; sin embargo pienso que lo más valioso de ella es la tabla misma. Se trata de un excelente compendio de las cualidades que los usuarios buscan en las revistas electrónicas.

Hoy en día las revistas electrónicas van perfilando su “personalidad propia” en la medida que la madurez intrínseca de los conceptos y la tecnología van aumentando. Actualmente, las revistas electrónicas van adquiriendo características, especializaciones, comportamientos cada vez más propios del medio electrónico. Gerry McKiernan de la Universidad Estatal de Iowa ha elaborado un compendio sobre las características y funciones que presentan las revistas electrónicas. [McKiernan, 2000]:

acceso a e-libros | artículos condensados | publicación expedita | servicio de alerta | lectura alternativa | presentación alternativa | anotaciones o comentarios a e-journals | inteligencia artificial | bibliografías | compra de libros en-línea | sitios de “chat” | indizado de citas | revistas “en paquete” | depuración colaborativa | código de computadora y programas de cómputo | memorias de congresos | hojeador conceptual | ligado conceptual | autoría cooperativa | compendios de e-journals | correcciones | manipulación de datos | acceso a bases de datos | bases de datos interconectadas | demostraciones | directorios | e-journals por disciplina | foros de discusión | tesis y disertaciones | compra de documentos | opciones de copia o descarga | artículos dinámicos | acceso a otras e-publicaciones | listas de discusión electrónicas | envío electrónico de manuscritos, arbitraje, y revisión | sistemas expertos | acceso extendido | e-journals federados | control de fuentes, formato, y despliegue | búsqueda e indización | e-journals inteligentes | tablas, gráficas y cartas interactivas | bases de conocimiento | megaservicios de alerta | modelado | búsqueda sobre varios editores | procesamiento en lenguaje natural | arbitraje abierto | e-journals complementarios | acceso parcial | patentes | conexión a colegas en red | e-journals personalizados | encuestas y revisiones | historia de prepublicaciones y versiones | preguntas y respuestas | e-journals reactivos | participación de lectores | ligas referenciales | recursos referenciales | entes relacionados | relevancia ponderada | e-journals de revisión | búsquedas salvadas | hojas de cálculo | estándares | bases de datos estadísticas | especialistas en la

materia | datos suplementarios | e-journals sinópticos | informes técnicos
| e-journals temáticos | mapas conceptuales de tema | servicio de traducción
| estadísticas de uso | conferencias virtuales | exposiciones virtuales
| e-journals | archiveros virtuales | marcado xml.

Cabe destacar aquí que esta lista no significa que todas las revistas posean todas y cada una de las características; es un compendio con la suma de los atributos y funciones observados en el universo de este tipo de publicaciones. McKiernan compiló también la serie de revistas que poseen cada una de las características señaladas. Y debe subrayarse también que algunas de las características descritas ya existían en revistas en papel y no fueron inventadas por las revistas electrónicas.

Debido a este conjunto de ventajas, las revistas digitales han ido ganando poco a poco terreno en la última década y posicionándose cada vez más en las colecciones de las bibliotecas, sobre todo las académicas.

En un estudio realizado a fines del año 2003 por la organización “*Public Communication Group*” —PCG— acerca de las tendencias observadas por los bibliotecarios respecto a sus colecciones de revistas académicas, se obtuvieron algunas conclusiones interesantes: [Public Communication Group, 2003].

El estudio se llevó a cabo por medio de una encuesta realizada a 100 bibliotecarios de norteamérica (E.U.A., Canadá), Latinoamérica (Brasil, México) y Europa (Inglaterra, Francia, Holanda, Bélgica, Italia, Alemania) para el periodo 2000-2003. He aquí el resumen de lo principal:

En cuanto al porcentaje de aumento en revistas digitales, 75% de los bibliotecarios de norteamérica afirmaron que incrementaron sus colecciones; 38% en latinoamérica y 82% en Europa.

Respecto al formato usado, en norteamérica 50% de bibliotecarios reportaron tener ambos formatos, 25% en proceso de cambio y 25% prefieren papel. En latinoamérica 88% tienen ambos formatos, 12% en etapa de cambio, 0% prefiere papel. En Europa, 61% reportó ambos formatos, 30% en cambio y 9% prefiere papel.

En lo tocante a la razón principal del cambio, en norteamérica 50% de los bibliotecarios expresó que el principal factor eran los usuarios o académicos, y que el otro 50% se debía al personal de la biblioteca. En Latinoamérica 25% reportó que el cambio se debía a usuarios o

académicos, 13% al personal bibliotecario y 62% a presiones presupuestales. En Europa, 68% reportó que el cambio se debía a usuarios y académicos, 32% al personal, y 0% al presupuesto.

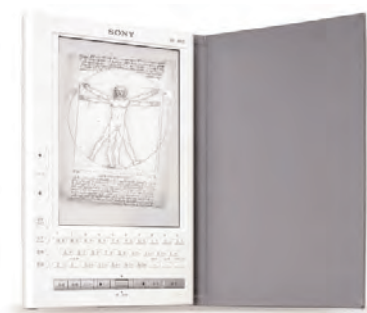
Como conclusiones finales se observó que:

- Las revistas digitales académicas van incrementándose cada vez más en las bibliotecas a nivel mundial a pesar de sus desventajas.
- Sólo una mínima proporción de bibliotecarios manifestaban seguir manteniendo colecciones de revistas exclusivamente en papel.
- Sólo 26% de los bibliotecarios estaban cancelando sus suscripciones en papel a cambio de la electrónica; 66% prefiere tener ambas.
- A nivel general, el motor principal del cambio a la versión electrónica proviene de los usuarios y académicos; en latinoamérica el motor principal son las restricciones presupuestales.
- El modelo de *Open Access* es medianamente popular en E.U.A. y Europa. No así en bibliotecas latinoamericanas.

3.3 LIBRO ELECTRÓNICO

“Con Internet, a los historiadores del libro les suele pasar como a los meteorólogos con las tormentas: antes que lamentar los efectos destructores, se muestran fascinados por los cambios y por las nuevas posibilidades que ofrecen”.

Robert Darnton.



El paleógrafo italiano Armando Petrucci nos dice que cada vez que una sociedad decide cambiar su soporte escriptorio por otro, lo hace respondiendo a presiones provenientes de las masas. El aumento del alfabetismo y la enorme necesidad de escribir y leer han provocado que una sociedad en la que abunda la producción de textos no pueda darse el lujo de tener

materiales de lectura caros.³⁹ Como muchas otras cosas en las bibliotecas digitales, la idea del libro electrónico no es producto de la era de la *Internet*; desde el punto de vista de ficción el famoso autor Isaac Asimov escribía ya en 1951 su cuento “¡Cómo se Divertían!”:

...Margie incluso lo escribió esa noche en su diario. En la página fechada Mayo 17 del 2157 ella asentó: «¡Hoy, Tommy encontró en el ático un libro de verdad!». Era un libro muy antiguo. El abuelo de Margie mencionó una vez que cuando él era pequeño su abuelo le había contado que hubo una época en que todas las historias venían impresas en papel. Los niños volvieron las páginas, que eran amarillentas y rugosas, y fue tremendamente divertido para ellos leer palabras que se mantenían quietas en vez de moverse como se suponía, sobre su pantalla. Y más aun, cuando volvían a las páginas anteriores, estaban allí las mismas palabras que ellos habían leído la primera vez...⁴⁰

Bajo el enfoque del mundo real, el libro electrónico data del año 1971, cuando Michael Hart obtuvo una “donación” considerable de tiempo de computadora en el equipo “Xerox Sigma V” en el Laboratorio de Investigación de Materiales de la Universidad de Illinois. Con ese recurso, él decidió que lo más valioso que podía integrarse a ese equipo era una colección de libros que pudiesen estar ahí almacenados y ser buscables y transferibles a otras computadoras sin pagos asociados. Por supuesto, esto sólo ha podido hacerse con obras que caen en el dominio público y que pueden ser transferidas bajo este contexto. El proyecto, denominado “*Gutenberg*” continúa hoy en día; recibe y distribuye obras digitalizadas por el público y a la fecha tiene alrededor de 7,000 títulos a disposición absoluta de los usuarios en la red y de manera gratuita.⁴¹

39 Petrucci, Armando. 1999. *Alfabetismo, Escritura y Sociedad*. Barcelona: Gedisa. p. 275.

40 Asimov, Isaac. 1974. The Fun They Had, en: *The Best of Isaac Asimov*. Doubleday. ISBN: 0-385-05078-X. Escrito originalmente en diciembre de 1951 para: “Boys and Girls Page” (syndicated newspaper feature) y publicado posteriormente en 1954 en la revista *The Magazine of Fantasy and Science Fiction* en febrero de 1954 disponible noviembre 3, 2005 en: <http://users.aber.ac.uk/dgc/funtheyhad.html> y en español en: <http://www.lorenzoservidor.com.ar/rel/rel321.htm>

41 Gutenberg Project. 2001. Página Web del Proyecto. Disponible noviembre 3, 2005 en: <http://promo.net/pg>

Como en los apartados anteriores conviene tratar de establecer la precisión del concepto *libro electrónico* de entre definiciones más generales de *publicación electrónica*. De acuerdo con esa antigua, simple, anacrónica y sin embargo todavía usada definición de UNESCO acerca de qué es un libro, sabemos que es *toda publicación impresa no periódica de al menos 49 páginas excluyendo cubiertas*. Difícilmente podríamos actualizar esta definición simplemente sustituyendo la palabra “impresa” por “electrónica” o “digital”. Tal vez en un principio y como ya hemos mencionado, antes los libros electrónicos han tendido a parecer una copia sobre una pantalla de lo que tradicionalmente era un libro impreso. Sin embargo, como otros medios, el libro electrónico comienza a desarrollar su propia personalidad y características —*con gracia*—.

Algunas enciclopedias en la red, como la *webopedia* dan definiciones tan irrelevantes como *la versión electrónica de un libro*, y luego continúan hablando de marcas comerciales de dispositivos de lectura. Según ODLIS se trata de:

un libro formado o escrito sobre una computadora, o convertido de letra impresa a formato digital —legible por computadora— proveniente de un proceso de escandido o similar, para su despliegue sobre una pantalla de computadora. Aunque la primer novela de hipertexto fuera publicada en 1987 —Afternoon, a Story por Michael Joyce—, los ‘e-libros’ no capturaron la atención del público hasta la publicación en-línea de la novela corta de Stephen King Riding the Bullet en el año 2000. [ODLIS, 2002]

Ana Arias Terry⁴² centra su definición en textos homólogos de papel:

en su nivel más simple, un libro electrónico consiste en el contenido electrónico originado de libros tradicionales, materiales de referencia, o revistas, puestos en la Internet y vistos con cualquier tipo de dispositivos, tales como: computadoras personales, portátiles o de palma, o lectores de e-libro dedicados.

⁴² Arias Terry, Ana. 1999. “Demystifying the e-Book - What is it, Where Will it Lead us, and Who’s in the Game?”. En: Against the Grain. Noviembre 1999. Disponible noviembre 3, 2005 en: http://www.against-the-grain.com/ATG_AnaEbook.html

La definición del “Proyecto de Libro Visual” —*Visual Book Project*—, se concentra en la preservación de la metáfora del familiar libro de papel:

el resultado de integrar la estructura de libro clásica, o más bien el familiar concepto de un libro, con rasgos que pueden ser proporcionados dentro de un ambiente electrónico, se menciona como un libro electrónico, que es interpretado como un documento interactivo que puede ser formado y leído sobre una computadora.

En palabras de Walt Crawford, “...cuando alguien te pregunte qué haces en torno a los libros electrónicos, lo mejor es preguntarle ¿qué entiendes tú por libro electrónico?”

Él mismo estableció una clasificación de libros electrónicos dividiéndolos en nueve categorías: [Crawford, 2000]

- *Dispositivos electrónicos propietarios e-book*: los dispositivos de equipamiento portátiles dedicados expresamente a la lectura tales como los de las marcas *Librié* de Sony, *Gemstar*, *Viewsonic*, *eBookman* de *Franklin*, HP, etcétera, en los cuales el texto es transferido en formatos propietarios típicos a ese tipo de lector.
- *e-books abiertos*: el estándar basado en XML que permite que cualquier texto resida sobre cualquier lector, protegiendo a la vez los intereses de los editores. Los potenciales dispositivos incluyen computadoras personales, portátiles, de mano, y los anteriormente mencionados dispositivos propietarios e-book.
- *Libros electrónicos gratuitos o libres*: copias digitales de libros ya del dominio público por haberse terminado la vigencia de sus derechos o porque así fueron entregados expresamente por sus autores, tales como aquellos disponibles a través de la *Biblioteca del Proyecto Gutenberg*. Éstos están libres para transferir, imprimir, copiar o circular.
- *Pseudolibros*: títulos comprados por bibliotecas o consorcios y prestados a usuarios externos para ser transferidos a sus propias computadoras personales. Solamente un usuario puede “tomar prestado” un título a la vez, a menos que la biblioteca haya pagado por más de una copia.

- “*Libros instantáneos*” o *Instabooks*: libros impresos y encuadernados bajo demanda expresa provenientes de textos digitales totalmente premarcados para impresión o imágenes de páginas escandidas y almacenadas digitalmente por compañías tales como *Lightning Source* o *Replica Books*. Estos libros se obtienen por medio de un pago o cuota.
- Los *no exactamente libro*: los textos de longitud media —novelas o novelas cortas— tales como la novela de Stephen King *Riding the Bullet*, que fueron creadas expresamente para ser distribuidas por la red, sin un antecedente impreso, e imitando la presentación física de una novela, normalmente como archivos PDF.
- *Autopublicaciones* o *e-vanidad (e-vanity)*: libros publicados personalmente por individuos en la *Web*.
- *e-libros “pre-web”*: libros publicados sobre CD-ROM o disquetes.
- *e-libros extendidos*: publicados sobre el CD-ROM o la *web*, éstos van más allá de las características típicas de los libros impresos con abundantes innovaciones, ofreciendo textos buscables, hipertextos, multimedios y elementos interactivos.

A propósito de esta última categoría, los “e-libros expandidos”, he aquí una lista más o menos consensada que obtuve de esas características inherentes hoy en día del “verdadero libro electrónico”, o “e-book”:

- Hipervínculos dentro del libro a contenidos *web* u otras herramientas de referencia, como diccionarios, atlas, catálogos, etcétera.
- Gráficas generadas en-línea de acuerdo a parámetros interactivos del usuario.
- Interacción con otros libros electrónicos.
- Compra autocontenida instantánea en-línea.
- Actualización automática o periódica del libro.
- Capacidad de información multimedial.
- Interacción con otros lectores.
- Traducción autocontenida.
- Búsqueda por palabras interna, subrayado, marcado, comentario, etcétera.
- Acceso a bases de datos o “bookmarks”.

Acerca de la primera categoría, los “dispositivos electrónicos propietarios *e-book*”, cabe aclarar que algunos de estos dispositivos de lectura han sido llamados *e-books* por sus fabricantes, lo cual ha creado más confusión en el medio, ya que no les permite a los usuarios establecer claramente cuál es el *e-book* o libro electrónico: ¿el soporte o el contenido? En un libro impreso son el mismo, pero conviene hacer la diferencia en un libro electrónico, ya que una cosa es el equipo para leer los libros y otra son los archivos con libros. Por ello algunos autores denominan ya *e-text* al contenido para evitar confusiones.

Joseph Esposito⁴³ distingue los libros impresos “tradicionales”, a los que denomina “libros originales” —*primal books*—, de los “libros procesados” o digitales —*processed books*—. Su división se basa en la manera de presentar el contenido, y no en su soporte; el “libro original” es el libro que conocemos: escrito por uno o pocos autores y visto como la materialización del pensamiento de uno o pocos individuos. El libro tratado, por otra parte, es el resultado de insertar plenamente al libro en un ambiente automatizado, conectado a una red. *Procesar* un libro significa bastante más que sólo construir vínculos con otro sitio; también incluye una modificación del acto de creación, que tiende a favorecer la integración de ese libro en una red de aplicaciones, que incluye comentarios, pero no se restringe a ellos. Tal libro típicamente tiene al menos cinco características: como texto autocontenido; es decir, la información que él mismo trae escrita; como portal electrónico o entrada a otras fuentes de información; como plataforma, en donde es usado como base de referencia por otros libros o teorías; como información preparada para ser procesada por máquinas, y como un nodo que interactúa con muchos otros en una red. Esto hará que cada vez más y más los libros electrónicos sean la participación colectiva de muchas personas.

¿Cómo funciona la industria del libro impreso hoy en día? para contestar esta pregunta, podemos utilizar los planteamientos básicos de

43 Esposito, Joseph. J. 2003. “*The Processed Book*”. En: First Monday. Disponible noviembre 3, 2005 en: http://www.firstmonday.dk/issues/issue8_3/esposito

Wade Roush⁴⁴ al respecto. Primero debe establecerse si el libro actual impreso en papel, no sirve ya más. De acuerdo con este autor, desde que Gutenberg imprimió la primera Biblia, los editores han estado perfeccionando cada vez más la industria del libro impreso y encuadernado, y han desarrollado una tecnología muy eficiente cuyo producto es sumamente fácil de producir y de usar. Los países desarrollados tienen una enorme infraestructura masiva para imprimir y comercializar libros en papel. Los autores y los editores tienen métodos más que probados para seleccionar, editar y comercializar libros, así como tratos contractuales bien establecidos respecto a la división de ganancias.

Si todo funciona tan bien en el mundo editorial de la imprenta, al menos a primera vista ¿por qué el afán de fijarlos usando una tecnología electrónica de fantasía?; una mirada entre bastidores revela que como portadores de información, los libros de papel sí tienen algunos inconvenientes serios. En primer lugar, tales libros están hechos de moléculas de celulosa, no de impulsos electrónicos, y esas moléculas tienen que venir de algún sitio. La fabricación de unos cuantos libros de papel significa la muerte de un árbol —lo cual reduce la capacidad del ecosistema global para generar el oxígeno y quitar el dióxido de carbono de la atmósfera—. El almacenaje físico de los libros y su transporte consume recursos adicionales. El proceso de venta en librería en sí mismo genera una basura tremenda, porque es difícil predecir el tamaño del mercado para un libro impreso, y los editores por lo general imprimen demasiadas copias y terminan picando aquellos que no son vendidos en un plazo razonable —las devoluciones comprenden en promedio el 40 % de los libros impresos—. Para el lector, conseguir un libro implica un viaje a la librería local o, si el libro fue comprado en línea, una espera de una a dos semanas, con ayuda de aviones y camiones contaminantes. Además, los libros son pesados, y ello hace difícil llevar más de unos cuantos a la vez. Y en muchas partes del mundo, los libros son tanto caros como difíciles de encontrar debido a la escasez de librerías y a una pobre infraestructura de distribución.

44 Roush, Wade. 2001. *"e-Book Basics: e-Book Primer"*. En: e-Book Web. Disponible noviembre 3, 2005 en: <http://www.ebookweb.org/basics/ebook.primer.htm>

Un libro electrónico, queriendo esto significar que el contenido de un libro ha sido convertido a impulsos electrónicos similares a los que contiene un documento de un procesador de palabras, no tiene ninguna de estas limitaciones. En esta forma digital un libro puede ser almacenado y transmitido con costos mínimos; el mayor costo implicado realmente está en escribirlo. Tanto desde una perspectiva económica como desde una ambiental, el libro electrónico parece una opción tan obvia que uno podría ser tentado a predecir que pronto todos los libros serán publicados electrónicamente. Pero hay varias características que trabajan en su contra y que hacen difícil todavía su advenimiento masivo. Algunas de ellas ya han sido mencionadas en el apartado de revistas electrónicas, pero no está de más extraer las características más perceptibles de las desventajas de los libros electrónicos, para tratar de calcular el tiempo y los requerimientos necesarios para su afianzamiento en el gusto de los lectores. Si bien estas características son un conjunto complejo, podemos distinguir de entre ellas las siguientes como las más sensibles:

Un gran nudo gordiano: el despliegue. Recordemos que *un libro no es simplemente la información que contiene; es también el soporte de esa información*. Si el libro electrónico debe tomar un lugar junto a y finalmente en lugar del libro de papel, necesitaremos que el equipo y los programas permitan hacer la lectura de éste tan cómoda y convenientemente como la lectura de un libro de papel, junto con otras ventajas irresistibles. La realidad actual es que ninguno de los dispositivos electrónicos en el mercado alcanza todavía ese estándar.

Ciertamente nuestra capacidad de convertir información en *bits* y transmitirla a través del mundo entero ha avanzado vertiginosamente, mucho más allá de nuestra capacidad de desplegarla de forma práctica, cómoda y económica. Aunque se habla ya de la posibilidad de que la industria desarrolle un verdadero “papel electrónico,” —es decir, *un material escribible delgado, flexible, barato, ecológico, con contraste y reusable*— y de una “tinta electrónica”, esta posibilidad se encuentra todavía a unos años de ser viable. Mientras tanto, la industria fabricante de equipos afronta un enorme reto al tratar de desplegar los textos de una página conjuntando tecnologías actuales, como los circuitos electrónicos, las pantallas de cuarzo o plasma, las baterías recargables de

cromo-níquel o litio, etcétera, en un afán de darle al lector el aspecto y la sensación de una hoja de papel impresa.

3.3.1 Dispositivos

Apunta Roger Chartier que estamos en la actualidad asistiendo a la «invención de un nuevo soporte del texto, *la pantalla*, otra forma de libro, pues puede hablarse también del *libro electrónico*». ⁴⁵ Pero ¿realmente se ha logrado este propósito? si observamos los dispositivos actuales para lectura de un libro notaremos que todavía adolecen de serias desventajas. Existen hasta la fecha tres tipos de equipos para esta tarea.

En primer lugar están las computadoras de escritorio, con su gran capacidad de almacenamiento y pantallas de nitidez y tamaño excelentes, pero tienen cero portabilidad; y no todo el mundo tiene la posibilidad de tener una en casa; y por otra parte su comodidad de lectura, a pesar de todas las emulaciones a un libro el resultado en la pantalla sigue siendo cuestionable. Según algunos estudios, las personas tienden a imprimir aquello que implique leer más de tres pantallas a la vez. Existen también las computadoras portátiles *laptop*, que resuelven el problema de la portabilidad a medias —no son absolutamente portátiles dado su peso y autonomía de baterías—, y siguen teniendo el problema del costo —son aún más caras que las de escritorio—.

En segundo lugar observamos que los ya mencionados dispositivos electrónicos propietarios creados específicamente para lectura e-book —“*Sony Data Discman Electronic Book Player*” y “*Sony Bookman*”; “*Franklin Digital Book System*”, “*Illumination Thoughtware*”, “*Gemstar*”, “*Viewsonic*”, “*HP*”, etcétera— han resuelto el problema del peso y la portabilidad y cuentan con excelentes pantallas, pero su capacidad de almacenamiento de documentos es realmente limitada, sus formatos han tendido a ser propietarios o exclusivos y su costo no ha entrado en rangos atractivos. De hecho, muchas de estas propuestas

⁴⁵ Chartier, Roger. 1999. “*Cultura Escrita, Literatura e Historia*”. México: Fondo de Cultura Económica. ISBN: 968-165974-0. pp. 48.

han sido sonados fracasos económicos. Las pantallas delgadas, ligeras y de calidad todavía están lejos de ser económicas. No obstante, la empresa *Sony* llega nuevamente con una propuesta en este sentido con su dispositivo “LIBRIÉ”, basado en el principio de la tinta electrónica, el cual parece ser ya una nueva generación en este tipo de dispositivos en cuanto a ligereza, nitidez, almacenamiento y precio. Habrá que esperar la opinión del mercado.

Finalmente, los dispositivos de lectura multipropósito manuales —computadoras de mano, PDA, *Handhelds* o *Palms*, y hasta teléfonos celulares— han tenido cierto éxito comercial con otros usos y han resuelto el problema de la portabilidad, pero sus capacidades de almacenamiento así como de batería, tamaño y calidad de imagen para lectura son francamente paupérrimas.

Como puede verse, cada tipo de dispositivo tiene algunas ventajas, pero también arrastra desventajas; por todo esto, hasta la fecha ningún dispositivo para lectura de libros electrónicos se acerca siquiera a la capacidad, comodidad, costo, portabilidad, confiabilidad, etcétera, que *en conjunto* tiene un libro.

Pero entonces ¿cómo debe ser el *dispositivo perfecto* para libros electrónicos? Debe ser delgado —*alrededor de un centímetro y medio de espesor*— y ligero —*no más de unos 400 gramos*—; tener una pantalla con el tamaño mínimo aproximado al “esquela” o a un libro de tapa dura. Dicha pantalla deberá ser cromática, legible inclusive a la luz del sol, tener una resolución sumamente alta —*al menos 200 puntos por pulgada*—, una proporción de contraste tan alta como la del papel, y debe poder leerse angularmente, es decir, de lado. El dispositivo deberá tener una capacidad de almacenamiento mínima de un Gigabyte —*mil millones de caracteres*—, bajo consumo de energía, y una duración de la batería muy larga —*medida en días, no en horas*— y por supuesto ser ésta recargable. El costo deberá estar en promedio por debajo de la barrera de los cien dólares.

Éste no es un reto fácil, el dispositivo así descrito requiere algunos de los elementos que hoy en día tienen la menor eficiencia por costo en la industria: los circuitos de memoria *ram*, las pantallas delgadas de cuarzo, cristal líquido o plasma y las baterías. Estos tres elementos han sido los que menos reducciones en costo han logrado y los incrementos

en capacidad son todavía muy onerosos, pero de todo ello depende mucho la aceptación final del libro electrónico. Robert Darnton,⁴⁶ en su obra acerca de la historia de la lectura, hace alusión a un grabado alemán de la Universidad de Leyden a principios del siglo XVII, que muestra:

los libros, innumerables volúmenes de pesados infolios, encadenados a altos estantes que sobresalen de las paredes. Los estudiantes leen de pie, en mostradores contruidos a la altura de los hombros, debajo de las estanterías; leen protegidos del frío por gruesas capas y sombreros, con un pie apoyado en una especie de taburete o saliente para aliviar la presión del cuerpo. Basta comparar ese grabado con los cuadros de lectura de fines del siglo XVII para ver cómo el desarrollo del confort crea espacios, tanto públicos como privados, que refuerzan el placer de encontrarse con los libros.

3.3.2 Seguridad

“Sit maledictus per Christum, Qui librum subtraxerit istum” -
“Que Cristo maldiga al que sustraiga este libro.”

El segundo elemento estratégico que ha requerido desarrollarse para el avance del libro electrónico es el aspecto de seguridad; es decir, debe evitarse que sea sustraído por manos no autorizadas. Y no hay nada nuevo en esta premisa, los libros y en general el material escrito eran en la antigüedad objetos preciosos. Requerían de personal especializado, largo tiempo de elaboración y materiales de soporte sofisticados. Por lo mismo eran caros de producir y delicados por naturaleza. Las bibliotecas que los han resguardado han tenido desde siempre el cuidado de preocuparse por su conservación, y han procurado preservarlos a toda costa; y no tan sólo por premisas económicas. Marc Drogin⁴⁷ afirma que *...cada civilización que creó para sí una forma de escritura, percibió ese don como proveniente de fuentes divinas;*

⁴⁶ Darnton, Robert. 1993. *Historia de la Lectura*, en: *Formas de Hacer Historia*. (Ed. Peter Burke), Madrid. p. 189.

⁴⁷ Drogin, Marc. 1983. *Anathema! Medieval Scribes and the History of Book Curses*. Totowa, NJ: Allenheld & Schram.

al escribir las cosas ellos estaban usando un instrumento divino y debían evitar que se abusara de él en cualquier forma. En una inscripción encontrada en una tabla del acervo de la biblioteca de un templo estatal en la Babilonia del rey de Nippur en el segundo milenio a.C. —recuérdese que los “libros” en ese entonces eran tabletas de arcilla con textos comerciales, religiosos, gubernamentales y educacionales— puede leerse esta advertencia, que sin duda puede considerarse como uno de los más antiguos dispositivos de seguridad bibliográfica.⁴⁸

Para la persona que sustraiga o quiebre esta tableta o la ponga en agua o la borre hasta que no pueda reconocerse y hasta su autor no pueda ni leer ni entenderla, ruego a los dioses Ashur, Sin, Shamash, Adad, Bel, Nergal, Ishtar de Nínive, Ishtar de Arbela, Ishtar de Bit Kidmurri, a todos los dioses del cielo y de esta tierra y a los dioses de Asiria, que maldigan a esta persona con una maldición que no pueda remediar, terrible y sin piedad, por todo el tiempo que él viva, y que su nombre muera, sus hijos y nietos sean exiliados y que su carne perezca en boca de los perros.

Por varios milenios, la manera de proteger la sustracción de obras escritas se basó en maldiciones y exhortos. Lawrence Thompson, en su estudio acerca de las bibliotecas medievales nos confirma que las maldiciones fueron también muy populares como método contra la sustracción de obras en el medioevo.⁴⁹ Como muestra de ello, en la colección de manuscritos del conde d'Angouleme, Jean d'Orleans, citada por Drogin,⁵⁰ puede verse esta maldición contra los sustractores de libros: *Que aquel que robe este libro, cuelgue en los cadalsos de París, y si no es colgado, que se ahogue, y si no se ahoga, que se queme, y si no se quema, que un fin todavía peor caiga sobre él.* Toda una serie de maldiciones y oprobios fueron arrojadas contra los sustractores de obras escritas. Las sanciones llegaron a incluir también al plagio de textos.

48 Crawford, Timothy. 1992. *Blessing and Curse in Syro-Palestinian Inscriptions of the Iron Age*. American University Studies: Series VII, Theology and Religion, Vol. 120. New York: Peter Lang Publishing. ISBN: 0820-41662-2. 259 p.

49 Thompson, James W. 1967. *The Medieval Library*. New York: Hafner Publishing.

50 Drogin, Marc. *Anathema!... Op. Cit.*

Adicionalmente se utilizaron excomuniones contra los infractores, como puede verse todavía hoy en día en la puerta de la Universidad de Salamanca o alguna vez se vio en las del Convento de San Francisco en la Ciudad de México; ahí se podía leer la carta del Papa Pío V, data-da en 1568, cuyo texto decía:

Para perpetua memoria. Según fuimos informados, algunos pródigos con su conciencia y enfermos de avaricia, no se avergüenzan de sacar, por gusto, los libros de antiguos monasterios y moradas de la Orden de San Francisco, y de retenerlos para su uso con peligro de sus almas y daño de las mismas bibliotecas, y no poca sospecha de los hermanos de esta orden [...] y condenamos a los sustrayentes a la sentencia de excomunión [...].⁵¹

De hecho, ya desde el tercer Concilio de Constantinopla del año 680 se decretó que cualquiera que sustrajese, vendiese o dañase un libro santo sería excomulgado. El arma final contra los infractores fue el anatema —*May the sword of anathema slay, If anyone steals this book away*—. La diferencia es que la excomunión sólo expulsaba al individuo de la eucaristía y otros sacramentos, mientras que el anatema incluía además una exclusión social y religiosa, y era incancelable salvo para unos cuantos altos prelados de la Iglesia. Significaba la exclusión y marginación total. Todos estos datos pueden estudiarse con más detenimiento en el estudio muy completo al efecto presentado en la obra *¡Anatema!* de Marc Drogin.⁵² Las maldiciones podían ser inclusive artísticas e ingeniosas. El mismo autor consigna este juego de palabras:

*Sor sup no scrip li poti
te er rum tor bri atur
Mor inf no rap li mori*

Si se lee combinando las sílabas del primer renglón con las del segundo, se leerá: *Sorte super norum scriptor libri poti* —“Que al escritor de este libro le sean dadas las dichas de la vida eterna”—. Y si se

51 Citado por: Torr e Villar, Ernesto de la. 1994. en *Ex Libris y Marcas de Fuego*. México: UNAM, Dirección General de Publicaciones y Fomento Editorial. ISBN: 968-36-7708-8. pp. 20-21.

52 Drogin, Marc. “Anathema!...Op. Cit.

combina el segundo con el tercero se leerá: *Morte infernorum raptor libri moriatur* —“Que el que robe este libro quede en las penas de la muerte infernal” —.

Otra forma de proteger las obras de la biblioteca de daño o sustracción era simplemente el evitar prestarlas. La siguiente regulación fue expedida por el Canciller de la Universidad de Oxford:

Dado que a lo largo del tiempo el gran número de estudiantes que usa la biblioteca es en muchos modos dañino para los libros [...] la universidad decreta que sólo los graduados y los miembros de las órdenes religiosas que hayan estudiado filosofía por ocho años podrán estudiar en la biblioteca [...] estas personas deberán prestar juramento de que no dañarán los libros [...].⁵³

Una tercera forma de proteger la sustracción de una obra era sin duda adicionarlas de ex-libris y marcas de fuego, las que las identificaban como propiedad de alguien; en palabras de Francisco Esteve Botey “*el signo de posesión o pertenencia, escrito o impreso, que desde tiempo inmemorial viene haciendo figurar todo dueño de un libro como señal de dominio o de propiedad particular*”.⁵⁴ Por supuesto, para abundar en este tema, está la obra clásica de Ernesto de la Torre Villar: *Ex Libris y Marcas de Fuego* [Torre Villar, 1994].

Lo anterior introduce la idea de que en algún momento las medidas de seguridad virtuales empezaron también a verse complementadas con medidas físicas. Al efecto podemos mencionar también la ampliamente conocida costumbre de encadenar los libros a sus estantes como medida de seguridad contra el robo, que estuvo en boga en los siglos XV y XVI.

Si ya en aquellos tiempos de pesados volúmenes las obras eran sustraídas, en estos tiempos de redes, computadoras, impresoras láser y libros virtuales nada más fácil que sacar una copia de un documento

53 Clark, J. W. 1902. *The Care of Books*. Cambridge: Cambridge University Press, p. 59. Citado por William Augustus Banner, profesor de filosofía de la Universidad Howards, en el discurso inaugural de la biblioteca de la universidad. Septiembre 23, 1983.

54 Esteve Botey, Francisco . 1949. “Ex-libris y ex-libristas”. En: *Madrid, Coleccionismo*, núm. 170 al 181, año XVI y XVII, 1929-30. Madrid, Aguilar.

etéreo. Una vez que un libro electrónico ha sido creado, ¿cómo evitar que rápidamente cientos, miles de copias no autorizadas circulen por la red? El libro de papel ha llegado a puntos de equilibrio con sus competidores. Salvo libros muy caros, o no muy comunes, no es práctico ni rentable fotocopiar íntegramente un libro, al margen de consideraciones de propiedad. Mucho menos hacer una copia mecanografiada. En general, el precio del libro impreso ha llegado a tal punto de equilibrio con los elementos que permiten copiarlo, que la mayoría de las veces no es atractivo para casi nadie hacer una copia de un libro en lugar del original.

No es el caso de un libro que está en forma electrónica; las maneras y soportes para copiarlo son numerosas, y los costos de ello son ínfimos. Cuando el propósito del editor es maximizar la distribución sin miras económicas, la tecnología permite hacerlo de manera muy eficaz y rentable desde hace tiempo. Cuando la edición tiene como propósito la venta y con ello una utilidad económica, el panorama cambia. Muchos editores se han detenido por mucho tiempo y se manejan con suma cautela ante este hecho antes de lanzar libros electrónicos al mercado, y no sin razón.

El 14 de marzo del 2000 Simon & Schuster lanzó al mercado la ya mencionada novela de Stephen King *"Riding the Bullet"*, primera en ser ofrecida a la venta sólo como libro electrónico, sin tener previamente una versión impresa en papel. Recibieron órdenes por más de 400,000 copias las primeras 24 horas y el libro se posicionó como el primer *"best-seller"* electrónico y por lo mismo marcó un hito en la historia de las publicaciones. Se afirma que King ganó 450,000 dólares con este libro. Cabe mencionar también que la compra en-línea de este libro generó varias decenas de miles de reclamaciones de los clientes a los pocos días, y rompió también *record* en este sentido. Igualmente, a las siguientes 24 horas de su aparición, ya podían encontrarse copias piratas de la obra en la red lo que de nuevo trae a colación la pregunta de cómo prevenir este tipo de actos. Varias técnicas han sido introducidas con este propósito. Examinando las más comunes distinguimos:

1) La “mar ca de agua” o “filigrana” electrónica. De acuerdo con Agustín Millares Carlo:⁵⁵

la marca de agua o filigrana eran dibujos realizados en el papel durante el proceso de fabricación. Se realizaban cuando la pulpa estaba aún húmeda mediante unos alambres que se entrecruzaban entre los corondeles — hilos verticales— y los puntizones —hilos horizontales—; estos dibujos podían verse a contraluz.

Inventada en la segunda mitad del siglo XIII en la fábrica de papel de Fabriano en La Toscana, Italia, hoy en día se sigue usando en el papel moneda, cheques o certificados, documentos de identidad, etcétera, con fines de comprobar la autenticidad del origen. Por extensión la marca de agua o, más propiamente dicho, la filigrana electrónica consiste en la adición de un código oculto embebido dentro de un documento electrónico —texto, imagen pieza de audio, video, etcétera—, que puede ser explícito para el usuario o no y que identifica a ese documento, con el fin de:

- Confirmar el derecho de propiedad sobre alguna obra digital.
- Garantizar el origen de ese documento.
- Introducir en cada copia de trabajo una huella digital electrónica.
- Establecer una prueba de compra de esa obra.
- Proteger el contenido digital.
- Identificar filigrana digital y comprobar la integridad del contenido.
- Marcar un trabajo digital, cuando la filigrana digital guarda la información adicional sobre el producto sí mismo.

Esta idea, como la técnica de tratamiento del papel que le da nombre, no es nueva. En 1954 Emil Hembrooke⁵⁶ de la *Muzac Corporation*

55 Millar es Carlo, Agustín. 1971. *Introducción a la Historia del Libro y de las Bibliotecas*, México: Fondo de Cultura Económica.

56 Citado por: Cox, Ingemar J. y Miller, Matt L. 2001. “Electronic Watermarking: The First 50 Years.” En: *Proceedings of the IEEE 2001 Int. Workshop on MultiMedia Signal Processing*, 2001. Disponible noviembre 3, 2005 en: <http://www.ee.ucl.ac.uk/~icox/papers/2001/mmssp01.pdf>

solicitó y obtuvo una patente titulada *Identification of sound and like signals* —Identificación del sonido y señales afines— en la cual se describe un método para insertar un código de identificación, imperceptible al oído humano, dentro de la música grabada, con fines de probar la propiedad de la misma. En la patente se describe ya este método con su nombre permanente: “una marca de agua electrónica”.

Esta técnica se conoce con el nombre de *esteganografía*; es decir, la técnica de ocultar un mensaje invisible dentro de uno visible. En la actualidad se basa principalmente en algoritmos y técnicas matemáticos. No debe ser confundida con la *criptografía*, que es la técnica de ocultar la información contenida en un mensaje. La esteganografía tiene como propósito encubrir la existencia misma del mensaje.

- 2) La “huella digital electrónica”. Yo la llamaría más bien el *ex-libris electrónico*. En cuanto a construcción funciona de manera semejante a la técnica anterior, pero con la diferencia que su propósito es agregarle una etiqueta o identificador único a un documento digital, distinguiéndolo entre todos sus semejantes, y que viaje siempre con él, al igual que lo hace una huella digital con un ser humano. Por lo general consiste en un número codificado que responde a un cierto patrón de identificación, y su principal función consiste en la autenticación de documentos y en el seguimiento o trazado de los “viajes” o copias que hace el propio documento.
- 3) Encriptado. Con esta técnica que, como ya se mencionó, forma parte de la criptografía, el archivo que contiene el documento electrónico se modifica para que no sea visible por los mecanismos normales de lectura y/o edición de textos de una computadora, o para que no puedan ser efectuadas las funciones típicas que se realizan con documentos, tales como impresión, corte y pegado, modificación, etcétera. Un ejemplo de esto son los archivos “PDF” de *Adobe*. Como se observa en este formato un archivo así editado no puede ser modificado, ni pueden “cortarse” o copiarse partes del mismo y, si así se establece previamente, no puede ser impreso.

Como todo documento encriptado, descriptarlo y volverlo útil depende de una “llave”. Esa llave es un número o cadena de caracteres que permite el acceso a la información. En algunos casos el documento no porta llave alguna y sólo permite realizar algunas funciones preestablecidas. En otros casos, la “llave” se liga al número del procesador o disco duro de una computadora, con el fin de que se “ate” a ese equipo y no funcione en ningún otro, si bien este mecanismo es altamente rechazado por los usuarios precisamente por su “anclaje” a un sólo equipo. En otros casos más la llave puede ser un dispositivo pequeño, económico, del tipo conocido como “memoria flash”—el cual parece un pequeño llavero que se inserta en el conector “USB” de la computadora— la que lee el número de la *llave* de este dispositivo y descrypta el archivo. Esto permite que el usuario tenga múltiples copias del documento en varios equipos, pero sólo se activará aquél que tenga la llave en un momento dado; es decir, sólo una copia puede ser legible a la vez. Por la portabilidad que le otorga a un documento y por la garantía de unicidad que lo asocia con él tiende a ser mucho más aceptado por usuarios y editores.

- 4) **Temporalidad.** Con esta técnica la entrega de un documento electrónico tiene una vigencia finita. El documento sólo será legible mientras dure su vigencia, la cual puede ser preestablecida en horas, días, meses, etcétera. Una vez que el plazo transcurre, el documento se vuelve ilegible dentro de la computadora y es necesario volver a conseguirlo desde su fuente si se desea seguir trabajando con él. Este mecanismo ha sido muy atractivo para las bibliotecas digitales, ya que permite establecer el concepto del “prestamo electrónico” de un libro de una manera espléndida: la biblioteca le envía el libro a un usuario con una vigencia finita; después de ese tiempo el usuario ya no puede leer el documento en su computadora, a menos que lo solicite otra vez, siempre con una temporalidad preestablecida y finita. Bajo este principio la biblioteca puede enviar libros con liberalidad o cualquier otro documento a sus usuarios, en el entendido de que no los está regalando; realmente los está prestando. Sin embargo esto no ha resultado atractivo para el público que adquiere documentos.

Independientemente de la técnica que se utilice para la seguridad, es importante establecer que los soportes y los formatos tienen una fuerte relación con los contenidos, lo cual es necesario ir resolviendo para tener un adecuado equilibrio de los mismos, y por lo tanto se requiere evolucionar con ellos.

Lawrence Lessig,⁵⁷ en su libro *The Future of Ideas: The Fate of the Commons in a Connected World* comenta que a principios de 1970, la empresa RCA experimentaba con una nueva tecnología para distribuir películas cinematográficas sobre cinta magnetofónica —lo que hoy vendríamos a llamar el video—. Los investigadores buscaban, al mismo tiempo, encontrar un medio para reproducir el celuloide con alta fidelidad como descubrir un modo de controlar el uso de la tecnología. Su objetivo era un método que pudiese restringir el uso de una película distribuida sobre el video y le permitiera al estudio maximizar el retorno económico por el uso de la película.

La tecnología eventualmente escogida para el segundo propósito era relativamente sencilla. Una cinta de video sería reproducida una vez y cuando llegase al final, el casete se trabaría en ese lugar. Si un cliente quisiera reproducir la cinta otra vez, tendría que volver al videoclub y una vez ahí los empleados lo destrabarían por una nueva cuota. De esta manera, la empresa que poseyera la película tendría garantizada una compensación por cada uso de su material registrado. La RCA presentó un prototipo de esta tecnología a la Compañía Disney a principios de 1974. Pat Feely, un joven investigador de tecnología, hizo la presentación del dispositivo a cinco ejecutivos de Disney, quienes quedaron horrorizados. Feely comentó que ellos afirmaron que *nunca* permitirían que sus materiales fueran distribuidos en esa forma, porque el contenido —*por confiable que el mecanismo de cerradura fuese*— nunca estaría controlado suficientemente. ¿Cómo podrían saber ellos —preguntó uno de los ejecutivos de Disney— cuántas personas estarían sentadas allí mirando esa película? ¿Cómo impedir que cualquiera llegase por ahí y se quedara a verla gratis?

57 Lessig, Lawrence. 2001. *The Future of Ideas: The Fate of the Commons in a Connected World*. New York : Random House. ISBN 0-375-50578-4.

3.3.3 Formatos

En cuanto a los formatos de edición de libros electrónicos, ¿cuáles son los más utilizados hoy en día?

De inicio, un libro electrónico puede editarse en cualquier formato de texto usando para ello cualquiera de los formatos típicos de este tipo de procesamiento: DOC, TXT, SXW, WPD, etcétera. También pueden editarse en “html”, que como ya se ha mencionado es un formato usado ampliamente para envío de información en la red; puede editarse en formatos “pictográficos” como imagen de páginas, tales como “JPG”, “PCX”, “GIF”, etcétera, o en su versión “texto-pictográfica” como es el PDF. Todas ellas funcionan para este propósito; no obstante, éstas no son versiones “profesionales” de acuerdo con los estándares establecidos hoy en día por la industria editorial.

Para tratar de definir cuál es el estándar aceptado hoy en día, es necesario referirnos al “*Open eBook Forum*” —OEB—. Ésta es un asociación formada hace unos años por las empresas e instituciones más estrechamente relacionadas con el mundo del libro electrónico para discutir los temas relativos al medio e ir logrando consensos acerca de sus características, definiciones, estándares y recomendaciones. Como resultado de sus trabajos, en 1999 publicaron la especificación OEBPS 1.0 —*Open eBook Forum Publications Structure Specification*— la cual contiene la primera especificación consensada internacionalmente sobre la estructura de publicación de un libro electrónico. En agosto de 2002 este organismo publicó la versión 1.2 de estas recomendaciones, la cual contiene entre otras mejoras: nueva funcionalidad en el área del control de presentación; mejoras al vocabulario de marcado básico —ahora un subconjunto puro de XHTML 1.1—; la compatibilidad con la versión anterior de esta recomendación ha sido especialmente cuidada. Agrega un soporte de CSS2 enormemente ampliado; ésta especificación —*Cascade Style Sheet nivel 2*— define hojas de estilo por medio de un lenguaje de estilo que le permite a autores y usuarios agregar estilo —fuentes, espaciado, tamaño, y señales auditivas— a documentos estructurados —como documentos HTML y aplicaciones XML—. De este modo se separa el estilo de presentación de los documentos del con-

tenido de los mismos, el CSS2 simplifica la edición y mantenimiento en sitios *web* [Open eBook Forum, 2002].

Esta especificación ha sido construida y está íntimamente ligada a las estructuras y principios de XML. Como ya se mencionó, esa estructura permite establecer especificaciones perfectas acerca de la presentación del documento, y además extensas especificaciones sobre contenidos. Es compatible con registros “*Books MARC Format*”, con “*Dublin Core*” y con Z39.50, por lo que sin duda será rápidamente un estándar de edición “profesional” de libros electrónicos en la red.

Como resultado de estas especificaciones existen ya dos empresas miembros del OEB que han lanzado sendos productos que cumplen con el tratamiento de la recomendación, y que se perfilan como los dos medios de mayor penetración a corto plazo dentro de la edición profesional de libros electrónicos. Estas empresas son “Adobe”, conocida ya ampliamente por su formato “PDF”, quien ahora ha lanzado al mercado su nuevo producto generador de libros electrónicos de acuerdo con la recomendación 1.2 y que se denomina “PDF Merchant”, legible ya por la última versión del ampliamente difundido “Adobe Reader”. Por su parte, Microsoft lanzó su producto denominado “Microsoft Reader” el cual tiene las mismas funciones de edición electrónica, es acorde con la recomendación, tiene extensiones para llevar el material fácilmente al mundo de las computadoras de palma y, dada la capacidad de penetración del mercado de esta empresa, tiene amplias probabilidades de convertirse rápidamente en un estándar *de facto*.

Las empresas “*NetLibrary*” y “*Media Quest*” son ejemplos de los nuevos “intermediarios” en *e-libros* en el mundo académico; ambas compañías proveen la comercialización de contenidos “reempacados” para títulos anteriormente publicados tradicionalmente por universidades y editoriales comerciales a través de “licenciamientos” con un número grande de editoriales. Ambas ofrecen regalías a esos editores por títulos de su fondo editorial o títulos agotados considerados como trabajos que por otra parte no generarían más nuevas ventas en papel. Ubicada en Boulder, Colorado, y recientemente comprada por OCLC, *NetLibrary* ofrece títulos por licencias de sitio negociadas con bibliotecas universitarias. El modelo de acceso imita la edición impresa, con

acceso permitido a un sólo lector a la vez. Ningún título puede copiarse a la computadora del usuario, pero las páginas pueden imprimirse. Y los registros bibliográficos pueden ser añadidos al catálogo electrónico de la biblioteca, y permitir así el acceso de usuarios al título usando el URL agregado en esa ficha bibliográfica. *Quest Media*, que está ubicada en Houston, Texas, ofrece títulos a través de suscripciones individuales, pero no manejan el esquema de licencia de sitio, por lo que, el acceso a todos los títulos está abierto a cualquier número de lectores simultáneamente. Un motor propietario de búsqueda permite que los suscriptores descubran, desplieguen y exploren los textos. Al igual que la otra empresa, ningún texto pueden ser copiado a una computadora local, pero las páginas pueden imprimirse.

Finalmente, hubiera parecido obvio que la oferta de libros electrónicos a las personas sería una evolución natural de la industria de publicación así como un impulso a la comunicación académica eficaz; sin embargo, los modelos de acceso al *e-libro* que proporcionen las mismas ventajas que los modelos de acceso al libro impreso tienen aún mucho camino por andar. La nueva tecnología no es suficiente en sí misma para crear un mercado para los libros electrónicos; la mayor parte de las cuestiones que harán el éxito o la ruina de ese mercado no son de índole tecnológica; están basadas en preocupaciones legales, sociales y económicas, y tienen profundas raíces históricas. En este sentido, varios autores han subrayado cómo la figura ancestral del libro en la cultura occidental se ha entrelazado con nociones de identidad individual, tratado absoluto y verdad de representación; es decir, los libros son individuales, autónomos, y dan una perspectiva de cierta realidad. Por ejemplo, para Foucault, el libro es la representación preeminente, la imagen por antonomasia del autor [Foucault, 1977]; para Otlet, el libro es a la vez una máquina de conocimiento, un organismo y un contenedor de “energía del conocimiento” [Otlet, 1934]. Para Derrida, la “*civilización del libro*” es la civilización de la verdad o del “discurso completo” —los logos teológicos— [Derrida, 1986]. Para Deleuze y Guattari, a menudo el libro se organiza bajo una “*línea maestra del significado*” que dirige su contenido —*master signifier*—, y que puede ser, entre otros, el título, un tema, o el propio nombre del autor [Deleuze y Guattari, 1987]. Nuestras prácticas culturales comunes equiparan y mezclan la obra

con el libro y el autor, como el hecho de referir el texto de un libro como escrito por Dios y por ende su autoridad, o comentar “lo que un libro dice” como sinónimo de lo que un autor dice. Umberto Eco lo hace en “*El Nombre de la Rosa*” a través de su personaje Adso: “[...] *De pronto comprendí que a menudo los libros hablan de libros, o sea que es casi como si hablasen entre sí [...]*”.⁵⁸ Esto sólo por mencionar algunos ejemplos de eso a lo que se están refiriendo los autores mencionados previamente. Tales caracterizaciones son histórica y culturalmente específicas. Por ejemplo, como señala Roger Chartier, la construcción de la identidad autoral como “*garantía de la coherencia del libro*” se incrementó a partir del siglo xv, por lo menos en lo relativo a los textos vernáculos. Y con el advenimiento de la impresión mecánica de libros ya no iba quedando duda, el propio nombre y peso del autor iba garantizando como verdad sus palabras, siendo éstas dispersadas sobre un escenario geográfico y social cada vez más amplio y sin contar con la función de autenticación de escribanos y del clero. [Chartier, 1994].

Es necesario también estudiar al usuario más profundamente y entender su comportamiento alrededor del nuevo medio; sólo así podrán reunirse los complejos elementos para crear el ambiente en el cual los *e-libros* encontrarán su lugar apropiado. Notemos que lo especial de los libros electrónicos, lo verdaderamente distintivo, más allá de las cualidades o desventajas mencionadas, es que están cambiando las formas de leer. Si en la actualidad la lectura es percibida como una práctica, sujeta a usos y costumbres atávicos, es decir, históricos, sociales, grupales e individuales, la esencia de esa práctica es la manera cómo finalmente el lector se enfrenta con el texto; y aunque es cierto que la idea misma de texto es el producto construido de los condicionantes mencionados, y especialmente de la disposiciones retóricas del texto creadas por *una comunidad de interpretación*, los libros electrónicos aportan muchas posibilidades nuevas a ese texto potencialmente “móvil”. Como ha apuntado Chartier “[...] *los autores no escriben libros, escriben textos que se transforman en objetos escritos, manuscritos, grabados, impresos [...]*”. Él afirma que la construcción formal del libro ayudó a organizar prácticas específicas de la lectura. Sin embargo, no

58 Eco, Umberto. 1980. “*El Nombre de la Rosa*”. Ed. Lumen. ISBN: 84-264-2622-0. p. 349.

solamente es verdad que la “*lectura es siempre una práctica incorporada a actos, a espacios, y a hábitos*”, sino que también “*la lectura y la escritura son organizadas por la metafísica histórica del libro*” [Chartier, 1994]; es decir, por las creencias establecidas que van prevaleciendo referentes a la naturaleza del libro. El libro ha desempeñado no sólo el papel de representar y de incorporar la verdad, sino que también ha actuado como metáfora para la organización de prácticas y de espacios sociales más amplios, tales como la construcción y el uso de bibliotecas.

Carmen Iglesias⁵⁹ cita también a Chartier y a su obra “*Cultura Escrita, Literatura e Historia*”⁶⁰ acerca de la historia de la lectura, y afirma:

Los cambios en la forma del libro afectan a la misma, pues es interesante recordar que antes y después de la imprenta, los libros mantienen la misma estructura, la que nos es familiar también a nosotros. Es decir, el paso de las páginas una tras otra, con la ventaja de poderlas correr hacia atrás y hacia adelante, o recorrer sus párrafos con los ojos hacia arriba o hacia abajo, cuantas veces queramos. Como se sabe, no siempre había sido así. El paso del rollo que conocía la antigüedad a la revolución del códice tuvo lugar a lo largo de los siglos II, III y IV d.C. Hay que imaginar la dificultad de una lectura que tenía que ir desenrollando por un lado y enrollando por otro según se leía, sin la facilidad de volver hacia atrás o recorrer líneas hacia adelante mas que muy despacio, sin poder acotar en los márgenes con la soltura que se puede hacer en el códice todavía manuscrito y luego en el libro impreso. Todo ello agravado por el tipo de soporte material que podía tener el rollo y que fue, como es sabido, muy variado y sufrió muchos cambios durante la antigüedad, según se experimentaba con distintos materiales procedentes de plantas o de pieles de animales (del papiro egipcio al pergamino helenístico), hasta llegar poco a poco a la mejora y abaratamiento paulatino de la industria del papel. Todos estos cambios complejos de técnicas por un lado y de formas del libro por otro, «quiere decir —señala Chartier— que hay una autonomía de las revoluciones de

59 Iglesias, Carmen. 2002. “*De Historia y de Literatura como Elementos de Ficción*”. Discurso al tomar el nombramiento como miembro de la Real Academia de la Lengua Española. Madrid, septiembre 30, 2002. Disponible noviembre 3, 2005 en: [http://www.rae.es/rae/gestores/gespub000001.nsf/\(voanexos\)/arch5838FB807DBF5039C1256C7E0045FD95/\\$FILE/ciglesias.htm](http://www.rae.es/rae/gestores/gespub000001.nsf/(voanexos)/arch5838FB807DBF5039C1256C7E0045FD95/$FILE/ciglesias.htm)

60 Chartier, Roger. 1999. “*Cultura Escrita, Literatura e Historia*”. México: Fondo de Cultura Económica. ISBN: 968-165974-0. pp. 48, 50.

las prácticas culturales y éstas no pueden deducirse simplemente de las transformaciones técnicas o formales [...] estas diversas líneas de transformaciones tienen sus propias razones o lógicas, y cada una plantea problemas porque no estamos frente a un saber absoluto, estable».

3.4 OTRAS PUBLICACIONES ELECTRÓNICAS

Hemos analizado ya las dos publicaciones electrónicas más importantes en el ámbito de una biblioteca digital: revistas y libros. Pero por supuesto, éstas no son en modo alguno las únicas publicaciones de este tipo que pueden incorporarse a las colecciones de estas bibliotecas hoy en día.

En efecto, dado que las bibliotecas contienen colecciones de documentos, y los “documentos” en la actualidad pueden ser una variedad inmensa de posibilidades electrónicas, la cantidad de publicaciones digitales que pueden distribuirse y consultarse vía estas bibliotecas es inmensa. Simplemente en la variedad: “libros electrónicos” existen ya en todas sus variantes digitales: textos escolares, novelas, teatro, instrucción programada, tesis, diccionarios, atlas, reportes, etcétera. En las revistas electrónicas, o mejor dicho, “publicaciones periódicas electrónicas”, debemos considerar además de las publicaciones primarias de todo tipo, —académicas, especializadas, científicas, diarios, etcétera— las publicaciones secundarias también en todos sus tipos —índices, resúmenes, directorios, *abstracts*, tablas de contenido, analíticas, etcétera—. En realidad, todos estos documentos evolucionan día a día en su camino digital, y sus características van cambiando con el tiempo. Ramiro Lafuente señala: “*el documento en ambientes digitales se adapta y modifica de acuerdo a las necesidades e imposiciones de las nuevas tecnologías, lo que ha motivado que sea visto como un objeto único que representa nuevas realidades, usos y aplicaciones*”.⁶¹

Si bien cada tipo de “documento digital” o, más propiamente dicho, “objeto digital” tiene sus características, interrelaciones, elementos, muy

61 Lafuente, Ramir o. 1999. *Biblioteca Digital y Orden Documental*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas. ISBN: 968-36-7075-X.

particulares, básicamente seguirán los patrones de los dos tipos de documentos que ya hemos revisado. Por ello, no es el caso hacer una revisión detallada de todas y cada una de las posibles publicaciones electrónicas que pueden formar parte de las colecciones de una biblioteca digital. Me limitaré a hacer una relación somera del tipo de publicaciones de este tipo que pueden contenerse en alguna biblioteca digital. Por supuesto, estas bibliotecas no contienen todos y cada uno de estos tipo de publicaciones electrónicas, pero ciertas bibliotecas tendrán alguna colección relacionada con un tipo de publicación. He aquí la lista:

- Imágenes digitales: como parte de colecciones de fotografías, mapas, pinturas, dibujos, diagramas, bocetos, carteles, manuscritos, planos, partituras, esculturas, viñetas, patentes, estampillas, etcétera.
- Audio digital: como parte de colecciones de música, entrevistas, ponencias, mesas redondas, debates, lenguas, poesía, audiolibros, teatro, etcétera.
- Video digital: como parte de colecciones de cine, video educativo, ponencias, mesas redondas, entrevistas, entretenimiento, periodismo, etcétera.
- Bancos de datos de todo tipo: científicos, económicos, meteorológicos, geográficos, geológicos, sanitarios, estadísticos, astronómicos, jurídicos, antropológicos, educativos, etcétera.
- Sitios *Web*: con fines de entretenimiento, de comercialización, informativos, educativos, de organizaciones, gubernamentales, de servicios, etcétera.

A guisa de ejemplo, he aquí una lista de los tipos de publicaciones digitales que ya se encuentran en servicio en una biblioteca muy completa, la Biblioteca Británica:

Libros y revistas de todo tipo; atlas, enciclopedias, diccionarios, partituras, mapas, periódicos, semanarios, manuscritos, catálogos, encuadernaciones, iluminaciones, documentos personales, impresos, dibujos y fotografías, manuscritos, archivos, filatelia, patentes, grabaciones digitales de música, naturaleza, narraciones, teatro, literatura, lenguas y dialectos.

Finalmente y como otro ejemplo, la lista de las publicaciones digitales que la Universidad de Cornell⁶² oficialmente considera importantes para desarrollar; está formada por los materiales en los que la universidad es reconocida académicamente y en los cuales sus académicos colaboran. En términos generales se encuentran:

- Libros, artículos, reportes y otras publicaciones de los académicos de la universidad.
- Revistas académicas que edita la universidad.
- Memorias de congresos, conferencias, etcétera, que se organicen en o por la universidad.
- Reimpresión de monografías publicadas por la universidad actualmente agotadas.
- Documentos digitales que incorporen elementos multimedia que no puedan ser entregados en papel.
- Tesis digitales.

3.5 HIPERVÍNCULOS CON OTRAS COLECCIONES.

Hemos mencionado ya que otra de las fuentes para desarrollar una colección digital son los hipervínculos que pueden establecerse con los materiales de otras bibliotecas digitales. En efecto, existen muchos documentos electrónicos en la actualidad pertenecientes a otras bibliotecas digitales y con libre acceso. En este caso, una de las nuevas funciones del bibliotecario en la actualidad es ir identificando esas ligas o vínculos que aportan material interesante para nuestros usuarios, con el propósito de que puedan ser vistos y acceder a ellos desde la página de nuestra biblioteca.

Así, hay materiales que realmente poseemos físicamente en nuestra biblioteca digital ya que poseemos esa información en el interior de nuestra computadora, en sus discos. Hay materiales que no poseemos

⁶² University of Cornell. Electronic Publishing Steering Committee. 1998. Electronic Publishing Strategies for Cornell University. Disponible noviembre 3, 2005 en: <http://www.library.cornell.edu/ulib/pubs/EPSCFinalReport1998.htm>

físicamente pero que mediante un convenio, acuerdo o estipendio podemos traer de un tercero en forma gratuita o pagada, como es el caso de nuestros convenios de intercambio de materiales y de la adquisición o suscripción por parte de proveedores. En estos casos, no poseemos físicamente la información en nuestras máquinas, pero sí el derecho de poder acceder a esos materiales, sin que físicamente los tengamos en nuestro poder. De la misma forma, hay mucho material de acceso libre en la red al cual para tener acceso no necesitamos de convenios ni pagos: basta con establecer el vínculo electrónico. De esta forma, por ejemplo, un bibliotecario que quiere desarrollar una colección de textos en bibliotecología puede digitalizar documentos de este tema que posea en su poder, siguiendo ciertos criterios, y ponerlos en su servidor; y de la misma forma puede integrar a su computadora materiales originales digitales que genera su organización acerca de este tema; y puede poner en su catálogo el material digital que adquiere por compra o suscripción a proveedores pero que físicamente no posee. En este mismo contexto, puede integrar a esa colección otros documentos a los que tiene acceso por medio de un convenio, y puede además establecer en su sitio *ligas* con otras colecciones interesantes en el tema y que son de acceso libre, como por ejemplo la revista *Investigación Bibliotecológica* del Centro Universitario de Investigaciones Bibliotecológicas —CUIB— de la Universidad Nacional Autónoma de México —UNAM—, o el JEP, *Journal of Electronic Publishing* de la Universidad de Michigan, las cuales son de acceso libre y gratuito para todo el público, y hacerle creer a sus propios usuarios que forman parte de su propia colección al aparecer en sus catálogos de revistas en el tema de la bibliotecología. No existe ninguna restricción legal en la *Internet* para hacer referencia a un sitio Web.

Finalmente, juntando todos estos elementos revisados anteriormente, el bibliotecario moderno puede ir creando una colección digital planeada, armónica, vasta, poderosa, actual, etcétera. Como pudo verse, y éste es el gran secreto escondido de las buenas bibliotecas digitales, las colecciones de estas bibliotecas no están formadas sóloamente con lo que poseemos físicamente en nuestras computadoras, sino con toda una combinación de materiales propios yexter nos que en resumen, contienen:

- Documentos originalmente en soportes tradicionales, digitalizados por nuestra biblioteca u organización: fondos antiguos, libros y revistas, material agotado, tesis retrospectivas, etcétera.
- Documentos originalmente digitales producidos por nuestra organización: libros, artículos, tesis, reportes, bancos de datos y otras publicaciones de la institución; revistas académicas, memorias de congresos, conferencias, etcétera, organizados por nuestra institución.
- Documentos adquiridos directamente a proveedores comerciales: libros y revistas electrónicos, bancos de datos, etcétera.
- Documentos adquiridos indirectamente a un proveedor comercial: tablas de contenido con servicios de documentación electrónica, etcétera.
- Documentos adquiridos o intercambiados gracias a convenios o acuerdos con otras instituciones.
- Hipervínculos con otros documentos y colecciones de interés para nuestra comunidad a los que se puede acceder libremente en la red.

Las colecciones así formadas potencian enormemente nuestra colección original digital, la que conformada así puede ser de enorme provecho para nuestros usuarios y alcanzar un mejor costo-beneficio que una colección basada sólo en digitalizar, o sólo en comprar, o sólo en establecer ligas con terceros. Esta mezcla es sin duda la mejor opción para una colección digital.

Capítulo 4

Herramientas para el registro de documentos digitales

*“Nos estamos abogando en información
y sin embargo estamos muriendo
de sed de conocimiento”.*

John Naisbitt

4.1 DEL REGISTRO DOCUMENTAL MANUAL AL ELECTRÓNICO

Tan importante como la creación y desarrollo de colecciones digitales es el desarrollo de mecanismos que permitan su recuperación y distribución. De nada sirve crear copiosas cantidades de información si ésta no puede ser localizada cuando se requiere o si se deben invertir enormes cantidades de tiempo en decantar información no relevante para nuestra actividad. La información para el estudio, la investigación, el trabajo o la capacitación debe estar a unos cuantos golpes de tecla cuando se requiera. Debe ser veraz, oportuna, pertinente y suficiente. De hecho, ésta es la diferencia esencial entre una verdadera colección y una masa amorfa de información.

Por supuesto, entre más rica y abundante es una colección, es más adecuada para atender las necesidades de un mayor número de comunidades; pero mientras más numerosa sea una colección, más crece el problema de recuperar esa información de manera adecuada. Las bibliotecas han contendido con el problema a nivel mundial desde hace más de cien años, y por lo mismo han desarrollado sistemas muy refinados de almacenamiento y recuperación de información.

Desde fines del siglo XIX y a lo largo del XX, la bibliotecología mundial fue creando y perfeccionando múltiples herramientas en pos del siempre deseado “Control Bibliográfico Universal” —CUB—. Como muestra de las grandes aportaciones de la centuria están las reglas angloamericanas de catalogación, los sistemas de clasificación decimal, LC, los *tesauri*, los índices KWIC y KWOC, los números ISBN e ISSN, los esquemas de materias, los catálogos, las autoridades, las tablas de transliteración, entre muchos otros ejemplos. Todo esto antes del advenimiento de las computadoras. A esta lista habrá que agregar las aportaciones de

la bibliotecología ya con las tecnologías de información: MARC, lenguajes de marcado, Dublin Core, metadatos, etcétera, la lista se va complementando.

Cuando se formalizó la idea del Control Bibliográfico Universal, hace poco más de tres décadas, la comunidad bibliotecaria internacional todavía no tenía una idea muy clara de la importancia que tenía la normalización internacional para la automatización bibliográfica, y este tipo de proyectos eran más un ideal que un proyecto práctico. La idea de que cada documento fuese catalogado sólo una vez en su país de origen y de que el registro resultante estuviese disponible para toda la comunidad mundial parecía estar muy lejos de la realidad. Los registros se intercambiaban ya entre países —principalmente entre bibliotecas nacionales— en papel, pero dado que eran el resultado de diferentes reglas y prácticas de catalogación, su integración en un único catálogo era muy difícil. Si se incorporaban los registros internacionales sin cambio alguno se degradaría el catálogo muy rápidamente; si se realizaba una revisión detallada y una recaptura el costo iba a ser prácticamente el de efectuarlos como originales. El formato MARC estaba en sus comienzos cuando se lanzó la idea del CBU, y todavía se estaba redactando la Descripción Bibliográfica Internacional Normalizada (ISBD). Las reglas de catalogación de diferentes países carecían de una base común para la asignación y forma de los puntos de acceso —tipo “encabezamiento”— y seguían diferentes normas de descripción. De acuerdo con Michael Gorman fue la confluencia de una necesidad —*las bibliotecas nacionales y académicas de todo el mundo necesitaban una catalogación más barata y más actual*— y un medio —*la automatización y más específicamente, el formato MARC*— lo que nos ha llevado más cerca del CBU de lo que nadie se hubiera imaginado hace treinta años. [Gorman, 2001]. Con el tiempo, es fácil ver una trayectoria inevitable que hizo que el formato MARC, las normas ISBD y las Reglas de Catalogación Angloamericanas (AACR2) fueran las piezas claves que nos llevaron en relativamente corto tiempo muy cerca del control bibliográfico. Cada una de estas tres normas mencionadas tenía objetivos originales que estaban bastante lejos del impacto que posteriormente tuvieron en la normalización internacional. El formato MARC apareció originalmente para facilitar la creación de fichas catalográficas de la Biblioteca

del Congreso conforme las bibliotecas las solicitaran. La ISBD surgió a partir de la *Standard Bibliographic Description* (SBD) —*Norma para la Descripción Bibliográfica*— elaborada por una comisión designada para tal fin como resultado del Congreso Internacional de la IFLA de Expertos en Catalogación (IMCE). A la SBD se la consideraba, entre otras cosas, como un modo de normalizar la *presentación* de los datos descriptivos de un documento de tal modo que pudiese traducirse automáticamente al formato MARC —*de ahí que tuviese su puntuación peculiar, mecánica y exacta*—. La segunda edición de las Reglas de Catalogación Angloamericanas AACR2 fue el resultado de décadas de esfuerzo para conseguir la uniformidad en la catalogación en el mundo anglosajón y especialmente para unificar las prácticas de la catalogación descriptiva británica y norteamericana.

Cada uno de estos tres estándares se transformaron y tuvieron un impacto mucho más allá de lo previsto en casi todo el mundo. Es interesante recordar cómo y por qué cada uno de ellos se desarrolló y extendió porque necesitamos entender que el mundo bibliográfico —al igual que el mundo real— está lleno de consecuencias imprevistas y los efectos de una normativa pueden ser mucho mayores de lo esperado.

Es un hecho indubitable que el acumulamiento de experiencias por parte de la bibliotecología, lo preciso de las metodologías y el rigor de los procedimientos han logrado mucho en el control documental. Pero la siempre creciente expansión del número de documentos por controlar y la enorme variedad de los mismos impone una búsqueda eterna de nuevas herramientas más potentes para contender con la tarea. Las computadoras ayudaron mucho en el procesamiento de esta información pero trajeron a su vez nuevas variedades de documentos, los documentos digitales o electrónicos, en cantidades tan numerosas que nuevamente incrementaban el problema en rangos impensables. *Para volver accesible la cantidad de información y artículos diarios en la prensa, en las revistas, para conservar los folletos, informes, reportes, para encontrar las materias dispersas en los libros, para hacer un todo homogéneo de esas masas incoherentes, son precisos procedimientos nuevos, muy distintos a los de la antigua bibliotecología, tal y como han sido aplicados.* Si alguien piensa que esta afirmación

ha sido hecha derivada del auge de los documentos digitales, está equivocado; la dijo Paul Otlet en su famoso *Tratado de la Documentación* en 1934 [Otlet, 1934, p. 6].

Ha habido siempre un problema de balance en el registro documental desde que éste se formalizó: los mejores registros documentales son aquellos hechos por profesionales de esta actividad, siguiendo precisos métodos, auxiliándose con extensas y poderosas herramientas documentales al efecto, pero consumen mucho tiempo y recursos humanos calificados; requieren insumos caros, y no hay manera de seguirle el paso a la producción documental. Por otra parte, se han desarrollado métodos de registro diseñados para un personal mucho menos entrenado o aun para máquinas, con unos cuantos elementos prácticos de registro documental: son económicos, rápidos y fáciles de hacer, y están al alcance de casi cualquier persona o institución; pero la calidad de los registros es baja y crea problemas después durante la recuperación. Esta otra aproximación al problema permite procesar mucha más información pero conlleva falta de precisión después en la recuperación y queda sin procesar mucho material de tal forma que resultará de muy difícil acceso. Las preguntas son por tanto: ¿se procesa bien y con calidad sólo una parte del material documental o se trata de procesar todo con poca calidad? ¿cuál material electrónico vale la pena registrar?

Obviamente ambas posturas han creado debates desde hace muchísimo tiempo, y existen defensores de la calidad sobre la cantidad, que saben que mucho se quedará fuera, y otros que defienden que es mejor que todos los materiales estén procesados, aunque sea pobremente, en lugar de dejarlos sin algún tipo de proceso.

La respuesta a este debate no es sencilla y hay que analizar muchos elementos y examinar los entornos, contextos, recursos y tiempos de quien se hace la pregunta. Existen posiciones fundamentalistas en ambos extremos y están también quienes tratan de buscar algo que optimice ambas soluciones y se ubique en algún punto medio.

En efecto, ha habido personas y organizaciones que han defendido la ortodoxia del registro documental por muchos años sólo para ver cómo sus mecanismos eran rebasados vertiginosamente por la variedad y abundancia documental, lo que ha resultado en algo poco útil

para los usuarios, y otros han pensado que no era necesaria ninguna regla de registro, o que éstas podían ser del todo laxas, y que con la “fuerza bruta” del procesamiento en computadora toda la información podría ser encontrada. Ahí quedan como prueba algunos de los primeros “buscadores” de Internet, que recuperaban —*y algunos todavía lo hacen*— un verdadero océano de basura documental en el que encontrar lo relevante era tarea imposible, y que resultar on poco útiles para los usuarios.

Las experiencias han ido enseñándonos que esas posturas extremas y fundamentalistas han producido el mismo resultado enunciado en el párrafo anterior: *fueron poco útiles para los usuarios*. Por ello se ha seguido buscando ese delicado balance entre calidad y precisión aunado a rapidez y economía. Algunas cosas apuntan ya como mejores soluciones al problema, aunque no son las soluciones definitivas. Hemos aprendido de los errores y las posturas soberbias. Al fin y al cabo *el buen juicio proviene de la experiencia; la experiencia proviene del mal juicio*. Hemos avanzado bastante, pero aún estamos muy lejos de la solución.

Gorman, por ejemplo, como escéptico que es de la Internet, señala:

Estamos, entonces, a punto de conseguir el Control Bibliográfico Universal para todos los materiales ‘tradicionales’ (es decir, los no electrónicos) usando un formato aceptado universalmente para el intercambio de datos bibliográficos, una norma aceptada universalmente para la descripción bibliográfica y unas reglas de catalogación casi universales que se están usando o están influyendo en las reglas de la mayoría de los países del mundo. ¿Existe alguna razón en principio por la cual no deberíamos aplicar este sistema de control bibliográfico que ya tenemos a los documentos electrónicos? La respuesta es ‘no’. ¿Hay alguna razón en la práctica por la cual esta tarea sea difícil de llevar a cabo? La respuesta es ‘sí’. Ya he hablado sobre los problemas que presentan los recursos electrónicos y el uso de los ‘metadatos’ para someterlos al control bibliográfico [...] Otros documentos electrónicos son bastante diferentes y, por lo tanto, no parece que se puedan adaptar inmediatamente a las estructuras existentes de control bibliográfico. Sin embargo, si nos detenemos a reflexionar sobre ello podemos ver que hay una similitud entre los documentos sin importar su formato. Los documentos electrónicos tienen títulos, fechas, textos e ilustraciones, ediciones, editores, relaciones con otros documentos (ya sean electrónicos o no), autores, contribuidores. Sabemos bien cómo tratar cada

uno de estos elementos bibliográficos, cómo registrarlos, cómo realizar el control del vocabulario y cómo crear registros MARC que se puedan integrar en los catálogos de la biblioteca. ¿Por qué entonces hay mucha gente que se preocupa tanto del control bibliográfico de los documentos electrónicos o defienden soluciones como, por ejemplo, los metadatos, los sistemas expertos y los buscadores avanzados como alternativas a la catalogación? Yo creo que hay varias respuestas a la pregunta (sin excluir la ignorancia como factor), pero la más importante es la que está relacionada estrechamente con las características de los documentos de Internet. Todos conocemos las características de una biblioteca bien organizada. Son la organización, la recuperabilidad, la autenticidad y la invariabilidad. Hay quienes afirman que los documentos y sitios electrónicos (colecciones de documentos electrónicos) difieren de los demás formatos que los seres humanos han usado para comunicar y preservar el conocimiento a lo largo de los siglos sólo en que son un nuevo tipo de documento. (Esto no es nuevo, recordemos la preocupación de las bibliotecas norteamericanas por los materiales audiovisuales en los 60 y los 70). Entonces como ahora, se pensaba que los materiales audiovisuales necesitaban reglas de catalogación especiales y diferentes, bibliotecarios especializados y la transformación de la biblioteca en un 'centro de recursos'. El tumulto decayó a medida que la gente entraba en razón y los materiales audiovisuales se integraron en sus colecciones y en las reglas de catalogación y, por suerte, hoy todavía tenemos la Biblioteca del Congreso y no el Centro de Recursos del Congreso). El apoyo más fuerte a esta idea de singularidad procede de la fugacidad y mutabilidad de los documentos electrónicos. Estas características, que cualquier bibliotecario verdadero deplora, son realmente el resultado lógico de la historia de la comunicación humana, cada formato produce más documentos que sus predecesores y dura mucho menos. Hacer muchas copias de mensajes en piedra requiere mucho tiempo, pero estos mensajes pueden leerse milenios más tarde". [Gorman, 2001].

Por el otro lado, existen las posturas de quienes pensaron que la *Internet* era la solución final a los problemas de información y documentación. También como ejemplo, cito algunas afirmaciones que compiló Doug Johnson al efecto:⁶³

63 Johnson, Doug. 2002. "Why Do We Need Libraries When We Have the Internet?". En: *Real Questions, Good Answers*, vol. 2 no. 1. Disponible noviembre 3, 2005 en: <http://www.doug-johnson.com/dougwri/internet.html>

- *Las bibliotecas son para aquellos que no pueden pagar un módem*
- *Comprar libros es invertir en tecnología obsoleta; toda la información que cualquiera necesite estará disponible dentro de poco gratuitamente en Internet.*
- *Las tarifas de telecomunicaciones tendrán que salir del presupuesto de revistas.*
- *Nuestra nueva escuela no necesitará biblioteca ya que todas las aulas estarán en red.*

Obviamente, los años de la red han demostrado que no es así de simple. Todavía faltan muchas bibliotecas que crear dentro de la red para que ésta empiece a funcionar. Se ha hecho un experimento que ilustra muy claramente esta aseveración. Búsquese un término muy popular en uno de los “buscadores” de información de la red, actualmente *muy mejorados* con respecto a sus antecesores de la década, siglo y milenio pasados. Si se busca el término, digamos *Lord of the Rings* en el buscador *Google* la respuesta indica que recuperó 10.4 millones de referencias —eso sí, en tan sólo 0.22 segundos—. Sabiendo de antemano que el texto completo de la obra no existe ahí en la red, ¿para qué quiero más de seis millones de referencias? Peor aún, la misma pregunta en el buscador *Yahoo* indica que recupera 17 millones de referencias en tan sólo 0.23 segundos. Seguramente ahí debe estar lo que busco, pero ¿dónde?

Si pregunto en español, el panorama mejora, aunque no sensiblemente. La búsqueda *Señor de los anillos* —*atención, debe saberse que en este buscador hay que buscar con “n” ya que con la “ñ” prácticamente pierde todo*— arroja tan sólo 875,000 referencias en *Yahoo* y 1 060,000 en *Google* —este último sí diferencia entre la “n” y la “ñ”—. Eso sí, para nuestro beneplácito y admiración, las búsquedas sólo tomaron 0.12 y 0.18 segundos, respectivamente. ¿Qué obtuvimos si sólo queríamos saber si el libro estaba en alguna biblioteca o tienda cercana? Nada.

Obviamente, en la vida real las preguntas siempre deben ser hechas en el lugar y contexto adecuado, y ello nos permite converger hacia una solución; de otra manera recuperaremos mucha información

irrelevante, pero el ejemplo sirve para ilustrar a lo que se puede llegar sin hacer antes un balance adecuado de las premisas.

Como esta obra tiene como centro las bibliotecas digitales, conviene enfocarse en los mecanismos de registro y recuperación documental de acervos de ese tipo, comenzando esta revisión a partir de las herramientas electrónicas; no obstante, siempre es bueno tener presente que el registro documental no comenzó con la llegada de las computadoras a las bibliotecas, y que es una historia que tiene muchas más lecciones de las cuales podemos aprender y obtener experiencia, y que comenzó mucho tiempo antes.

4.2 METADATOS

Los metadatos en sí tampoco suponen algo muy nuevo dentro del mundo bibliotecario. Por mucho tiempo se afirmó que el primer catálogo en la historia de las bibliotecas se remonta veintitrés siglos atrás, y fue el registro temático en 120 volúmenes de la Biblioteca de Alejandría, denominado *Pinakes —Tablas—*, a cargo de su director Calímaco de Cirene, realizado cerca del año 265 a.C. Abundemos al respecto: como concepto, los metadatos existen en las bibliotecas desde que el primer encargado de una de ellas hizo un *listado* de las tabletas que se encontraban en su colección, hace ya cerca de cinco mil años. La expedición arqueológica italiana encabezada por Paolo Matthiae excavó en 1975 una colina llamada Tel Mardikh, cerca de Aleppo, en el noroeste de Siria, y descubrió, entre otras cosas, la biblioteca de Ebla dentro del palacio del Rey Aghrish. Ebla, *la Ciudad de las Piedras Blancas*, llegó a tener 260,000 habitantes, floreció en la Edad de Bronce temprana y fue destruida alrededor del año 2,300 a.C. Su biblioteca contenía más de 17,000 documentos sobre tabletas de arcilla que fueron datados a la mitad del tercer milenio a.C., por lo que muchos la consideran la biblioteca más antigua del mundo.⁶⁴ Algunas de las tablillas contenían

⁶⁴ Matthiae, Paolo. 1980. *Ebla: an Empire Rediscovered*. Trans. by Christopher Holme. Garden City, N.Y.: Doubleday. Citado por: Wellisch, Hans. 1981. "Ebla: the World's Oldest Library". En: *Journal of Library History*. vol. 16, no.3 (Summer 1981). p. 488-499.

un “listado” de las obras que se encontraban en la colección y por tanto se consideran, hasta ahora, como el metadato más antiguo que haya sido descubierto y datado; las tabletas habían sido originalmente ordenadas en sus anaqueles por temas; todo ello veinte siglos antes de la Biblioteca de Alejandría. Metadato es también la famosa *Bibliotheca Universalis* de Konrad de 1545;⁶⁵ la idea de crear una bibliografía universal animó al naturalista y escritor suizo-germano a compilar su obra, en la que recogió todos los libros impresos en Europa en latín, griego y hebreo. Tiene más de 12,000 referencias sobre unos 3,000 autores, clasificados por nombres. Se trata de la primera bibliografía general, seguida en 1548 de una tabla sistemática que distribuía los libros en veintiún secciones, y de otro apéndice en 1555. Sin duda esto ha sido un gran metadato por casi cinco siglos. El término *metadato* como tal fue acuñado por Jack Myers en 1969 para describir conjuntos de datos en una línea de productos de algo que él llamó su *metamodelo*. Posteriormente fundó una empresa con ese nombre —*Metadata*—. La primera acepción que se le dio al término —y actualmente la más extendida— fue la de *dato acerca del dato*, ya que proporcionaba la información mínima necesaria para identificar un recurso documental. En ese mismo sentido se afirma que un metadato puede incluir información descriptiva sobre el contexto, calidad y condición o características del dato. La evolución detallada del término desde esa fecha hasta 1997 ha sido descrita por Lange y Winkler.⁶⁶

En términos generales, un metadato es un elemento que describe el contenido, condiciones, características, etcétera, de un documento, con el fin de definir, identificar, organizar, indizar, filtrar, colocar, preservar, recuperar y administrar ese documento como una parte de un conjunto ordenado de recursos de información, electrónicos o no. Los metadatos pueden crearse para describir los atributos inherentes de un

65 Gesner , Konrad, 1545. “*Bibliotheca Universalis, Sive, Catalogus Omnium Scriptorum locupletissimus, in Tribus Linguis, Latina, Graeca, & Hebraica ...*”. Zürich. Apud Christoph Froschauer.

66 Lange, Holley y Winkler , B. Jean. 1997. “Taming the Internet: Metadata, a Work in Progress”. En: *Advances in Librarianship*, ed. Irene Godden, v. 21, p. 47-72. San Diego: Academic Press.

recurso documental cualquiera que éste sea: objeto bibliográfico, —*libro, revista, tesis, etcétera*—, registros e inventarios archivísticos, objetos geoespaciales, recursos visuales —*realia*— y los de museos y galerías, programas de computadora, etcétera. Los metadatos pueden presentar diferentes niveles de especificidad, profundidad o estructura.

Obviamente de acuerdo con la definición antes mencionada, podemos considerar a la catalogación como un proceso de generación de metadatos y a la ficha catalográfica como su metadato resultante. Teniendo en cuenta que la mayoría de sistemas de metadatos ha sido creada no sólo por profesionales bibliotecarios de la información, sino también por ingenieros en sistemas informáticos, archivistas, etcétera, la utilización de este término puede implicar una amplitud excesiva. Por ello, es mejor utilizar al metadato como un término neutral, de acuerdo con Caplan,⁶⁷ que permita alejarnos de posibles prejuicios por parte de todas aquellas personas menos cercanas al mundo bibliotecario, y que coloque a todos los grupos profesionales implicados en su desarrollo en una posición de igualdad.

El hecho importante que nos ocupa en este momento es que una buena parte de las colecciones de documentos en formatos tradicionales ya tiene algún tipo de “metadatos”: tarjetas catalográficas, registros MARC, tesauri, KWIC y KWOC, números ISBN o ISSN, etcétera. Si éstos pueden perfeccionarse, como Gorman afirma, estamos en camino de alcanzar el Control Bibliográfico Universal. El problema comienza al percatarnos de que, al revés de lo que sucede con los documentos en las bibliotecas, la inmensa mayoría de los recursos electrónicos existentes en la red no tienen un mecanismo de registro ordenado, un “metadato” que ayude a su registro y recuperación, y esto es motivo de una gran preocupación, ya que la simple extensión de los mecanismos y herramientas para los documentos “tradicionales” no basta, por el tipo de documentos electrónicos disponibles y por el número tan abundante de ellos. Por lo tanto, es de suma importancia establecer los avances y las posibilidades de los metadatos orientados hacia los

67 Caplan, Priscilla. 1995. “You Call it Corn, We Call it Syntax-independent Metadata for Document-like Objects”. En: *The Public Access Computer Systems Review*, v. 4, no. 6.

documentos electrónicos para las bibliotecas digitales. Si se analiza desde el punto de vista de la información distribuida, el metadato, como concepto, puede aportar información que va más allá del término catalogación.

Inciendiando sobre el aspecto básico de la definición, otros autores, amplían el concepto de *dato sobre el dato* al afirmar que incluyen información sobre su contexto, contenido y control, así como sobre todo lo que tenga que ver con “el dato”. Ercegovac,⁶⁸ por su parte, afirma que un metadato describe los atributos de un recurso, teniendo en cuenta que ese recurso puede consistir en un objeto bibliográfico, registros e inventarios archivísticos, objetos geoespaciales, recursos visuales y de museos o programas de computadora. Aunque los metadatos puedan presentar diferentes niveles de especificidad o estructura, el objetivo principal es el mismo: describir, identificar y definir un recurso para recuperar, filtrar, informar sobre condiciones de uso, autenticación y evaluación, preservación e interoperatividad. Murtha Baca indica algunas de las características relevantes de los metadatos: [Baca, 2000].

- Los metadatos no tienen que ser digitales por fuerza. Los profesionales de la información han estado creando conceptos y herramientas de acceso a la información durante milenios. Actualmente, los metadatos se agregan en forma digital a los sistemas de información.
- Los metadatos tienen que ver con algo más que la simple descripción de un objeto. Los profesionales de información: bibliotecas, archivos, museos, etcétera, están familiarizados con los conceptos de catalogación o la descripción de objetos documentales, pero los metadatos pueden usarse también para indicar contexto, administración, procesamiento, preservación, disposiciones legales, etcétera, del recurso que se describe.

68 Ercegovac, Zorana. 1999. “Introduction” within the Special Topic Issue, “Integrating Multiple Overlapping Metadata Standards”. En: *Journal of the American Society for Information Science*, v. 50, n. 13, p. 1165-1168.

- Los metadatos pueden venir de una gran variedad de fuentes. Pueden ser creados por una persona (autor, catalogador, usuario, etcétera), crearse automáticamente por una computadora, o ser inferidos a través de una relación con otro recurso, tal como un hipervínculo.
- Los metadatos tienden a incrementarse durante la vida de un objeto de información, y por tanto pueden ser creados, modificados, aumentados, etcétera, durante la vida del recurso.
- El metadato de un objeto documental puede ser simultáneamente el contenido de otro objeto documental. Una bibliografía, por ejemplo, es a la vez metadato y obra.

Senso y de la Rosa [2003] hacen un resumen de la mayoría de las funciones descritas por estos autores respecto a los metadatos helos aquí:

- Resumir el significado de los datos
- Permitir la búsqueda
- Determinar si el dato es el que se necesita
- Prevenir ciertos usos
- Recuperar y usar una copia del dato
- Mostrar instrucciones de cómo interpretar un dato
- Obtener información sobre los derechos de autor y uso
- Aportar información acerca de la vigencia del dato
- Ofrecer información relativa al propietario o creador
- Indicar relaciones con otros recursos
- Controlar la gestión documental

En estas definiciones podemos observar un salto cualitativo importante pues se considera que los documentos —así como sus partes: líneas, párrafos, imágenes, etcétera— se pueden tratar como objetos, y los metadatos como los atributos que definen las características de cada uno de ellos, sin limitarse a su descripción simple (lo que hasta ahora venía siendo la catalogación), y que nos servirán como base para analizar las herramientas más notables realacionadas con el registro de materiales electrónicos para las bibliotecas digitales.

Como resumen, podemos establecer tres tipos fundamentales de metadatos: *administrativos* (como el número del título o de inventario), *descriptivos* (como el autor, título, editorial, etcétera) y *técnicos* (como las restricciones legales, el tamaño del archivo, etcétera). Distinguimos también cuatro tipos de metadatos en relación con sus contenidos:

- 1) Metadatos separados físicamente del contenido del objeto; por ejemplo un libro más su tarjeta catalográfica o un libro y su registro MARC.
- 2) Metadatos embebidos en el contenido del objeto, como por ejemplo el *header* o encabezado dentro de un documento HTML.
- 3) Metadatos ligados al contenido del objeto, como por ejemplo un registro MARC con un vínculo a un URL.
- 4) Metadatos embebidos y ligados respecto al contenido del objeto, como por ejemplo un registro XML con un documento PDF.

4.3 FORMATOS DE MÁQUINA Y LENGUAJES DE MARCADO

4.3.1 MARC

Desde la década de los sesenta, el personal técnico relacionado con el intercambio de documentos a través de computadoras comenzó a preocuparse por estructurar esa información en una forma normalizada que facilitara el intercambio y la manipulación de esos materiales. En el ámbito de las bibliotecas surgió el proyecto del formato “MARC”, —*MAchine Readable Cataloguing*—, pionero en esta temática. Durante 1965, la Biblioteca del Congreso de los Estados Unidos (LC) estableció un proyecto piloto denominado “MARC I” cuyo objetivo era definir una metodología para crear registros catalográficos en un formato legible por una computadora. Un proyecto similar apareció en el Reino Unido en esa misma época, ordenado por el Consejo de la Biblioteca Británica —*British Library*—, con el fin de establecer un formato para registrar y explotar los registros que serían usados en la bibliografía británica. A este proyecto se le llamó “BNB MARC” —*British*

National Bibliography with MACHine Readable Cataloguing—. Ambos proyectos derivaron en poco tiempo en una cooperación en el ambiente de las bibliotecas anglosajonas y dieron origen en 1968 al proyecto “MARC II”. El resultado fue el formato ampliamente conocido para intercambiar registros catalográficos de monografías vía electrónica el cual fue complementado poco después con los consecuentes formatos para publicaciones periódicas, mapas, discos, etcétera, hasta cubrir prácticamente todo tipo de documentos existentes en las bibliotecas.

Con el tiempo, MARC fue tan exitoso y tan ampliamente utilizado en todo el mundo que empezaron a usarse múltiples variantes del mismo, hasta llegar a unas veinte, las que fueron adaptadas a las nuevas necesidades de otras comunidades en el planeta. Para 1977 se intentó definir un formato universal “UNIMARC” como una especie de “esperanto” para MARC con objeto de que los distintos “MARC” pudieran intercambiar registros entre sí usando UNIMARC como intermediario. Múltiples versiones perfeccionadas de este formato lo colocan en esa posición de ser el intercambio universal más aceptado a la fecha. Además el formato también empezó a ser definido por otras entidades de datos por ya sobrepasar la simple producción de catálogos, como los registros de autoridad, pues su diseño permitía que fuera usado en prácticamente cualquier plataforma de cómputo del mundo. Debido a su amplia aceptación MARC se convirtió en norma ANSI para la unión americana en 1971 (Z39.2 - 1971) y en norma ISO internacional en 1973 (ISO 2709 - 1973) (E). Entre las grandes aportaciones de MARC a la tecnología documental están las tablas y/o criterios para normalizar —idiomas, países, alfabetos, etcétera—.

En esencia, MARC no ha sido nunca un catálogo ni un método para catalogar, como muchos han pensado a lo largo del tiempo. La **gran** aportación de MARC al manejo documental, y que ha sido seguida por muchos formatos posteriormente hasta nuestros días, fue la de *marcar*; esto es, asociar una “etiqueta” a cada uno de los elementos que conforman una ficha catalográfica, con objeto de que las computadoras puedan identificar y catalogar sus partes y posteriormente efectuar múltiples acciones con ellas en beneficio de las personas. Este principio tan simple y tan obvio en nuestra época, marcó un hito en la his-

toria del manejo documental por medio de computadoras. Hoy ese formato se ha sofisticado y diversificado enormemente, pero el brillante principio se mantiene hasta nuestros días. Gorman afirma: *Desde cualquier punto de vista, el formato MARC es un logro histórico y ha sido el principal motor para la normalización internacional desde un punto de vista práctico. Literalmente, es el motor que ha hecho posible el control bibliográfico universal. La transformación de la catalogación ha sido larga y, en gran medida, satisfactoria.* [Gorman, 2001]

El concepto de asociar cada elemento de una ficha, y posteriormente de cualquier otro documento, con una “etiqueta” —llamada en inglés *label*, *tag* o *token*— fue conociéndose como “marcado” del documento, y ha sido de tal trascendencia que todos los formatos que han aparecido posteriormente para el registro de documentos se basan en el concepto del marcado, y así lo ostentan en su nombre —*markup* o *mark up*—. Abundaré más adelante en este concepto. Por lo anterior MARC puede ser considerado como el “abuelo” de todos los lenguajes de marcado documental para computadora que existen hoy [What is MARC, 1996]. Actualmente, la versión más actualizada del formato es la conocida como “*MARC 21*”, la cual establece en esencia cinco estándares: Datos Bibliográficos, Datos de Autoridad, Datos de Acervos, Datos de Clasificación e Información para la Comunidad. Incluye ya también múltiples características derivadas de sus sucesores, los nuevos “lenguajes de marcado”, entre ellos las de “XML” y “MODS”, los cuales revisaremos más adelante.

Fuera de las bibliotecas, a nivel general, y durante esa misma época de fines de los años sesenta, IBM creó “GML” —*Generalized Markup Language*— o “Lenguaje de marcado generalizado” para contender con las necesidades de sus propios sistemas internos de publicación. Luego un grupo formado por Goldfarb,⁶⁹ Mosher y Lorie estableció la primera especificación en 1969 para que esta empresa produjera de manera normalizada libros, reportes, manuales y otros tipos de documentos a partir de un único conjunto de archivos originales en una computadora.

69 Goldfarb, Charles F. 1996. *The Roots of SGML. A Personal Recollection*. Disponible noviembre 3, 2005 en: <http://www.sgmlsource.com/history/roots.htm>

Este grupo fue el primero en introducir conceptos como *Estructura jerárquica de documentos* y *Descriptores de tipos de documentos*. Muchas otras soluciones de propósito específico fueron introducidas por otras organizaciones para estructurar información documental, pero nada realmente que abarcara tal tarea a gran escala. GML es además el antepasado común de los procesadores de texto, tan utilizados actualmente. En 1978, el *American National Standards Institute* (ANSI) comenzó un proyecto tendiente a desarrollar un marcado de documentos que sirviera como estándar de ese país para el intercambio de información. El doctor Goldfarb se unió al grupo ANSI, que publicó en 1980 el primer borrador que ya para 1983 se convirtió en estándar de la industria con la clave GCA-101-1983.

La primera tecnología de estructuración de información documental normalizada de cierta significancia fue el “SGML” —*Standard Generalized Markup Language*— o “Lenguaje de marcado generalizado estándar”, el cual también provino de la empresa IBM a principios de los ochenta con toda la experiencia adquirida con el GML. Este lenguaje fue creado originalmente con el fin de formatear y organizar la documentación técnica y legal dentro de esa empresa, pero posteriormente fue expandido y adaptado para ser usado en una amplia variedad de empresas como un estándar para manejar todo tipo de información. En 1984, la International Organization for Standardization (ISO) decidió integrar un grupo para desarrollar una versión internacional del estándar y para 1986, cuando fue publicado, se convirtió en una norma ISO (ISO 8879:1986) *Information processing—Text and office systems—Standard Generalized Markup Language (SGML)*. La versión completa del estándar puede verse en el sitio oficial de la ISO en: <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=16387>

Si bien SGML es extremadamente poderoso, es igualmente complejo y requiere de una considerable cantidad de programación adicional para procesarlo. Debido a su complejidad y los recursos extras que requiere, SGML no era una opción viable para representar textos e hipertextos en las primeras épocas de la Internet, dadas las limitaciones en la capacidad de los equipos de esa época.

En 1989, Tim Berners-Lee y Anders Berglund, dos investigadores del Laboratorio Europeo de Partículas Físicas (CERN), crearon un lenguaje basado en etiquetas para marcar documentos técnicos y hacerlos así compartibles en la Internet. Este lenguaje fue expandido en 1990 a una versión simplificada del SGML llamada HTML —*Hyper-Text Markup Language*— “Lenguaje de marcado de hipertextos” y desde 1992 ha llegado a ser el estándar para el manejo de información en la *Web* [Johnson, M., 1999].

HTML pudo representar perfectamente información *estática* en una página *Web*; es decir, textos previamente establecidos, imágenes, botones, ligas, etcétera, pero pronto hubo necesidad de ligar las páginas a bases de datos con el fin de traer a las pantallas catálogos, estados de cuenta, textos completos, etcétera; es decir, información *dinámica* para la cual HTML no fue diseñado. Un conjunto de elementos añadidos surgió para HTML en los últimos años para resolver esta problemática, algunos de ellos muy exitosos, como los derivados del lenguaje Java, y otros que han caído ya en el olvido. Al final, HTML comienza a estar demasiado *remendado* y empieza a ser demasiado complejo para llenar los requerimientos del manejo de información de la actualidad.

Como resultado de esto muchos técnicos volvieron sus ojos nuevamente hacia SGML, ya que su complejidad era de origen y no resultado de adiciones, y ha tenido capacidad para manejar documentos complejos de variados tipos, y por otra parte las capacidades de equipos y redes se han incrementado sensiblemente en los últimos años. Otros grupos de personas empezaron a reescribir una versión simplificada de SGML pero capaz de contender con las carencias de HTML, y de ahí surgió un nuevo lenguaje llamado XML: el *eXtensible Markup Language* o “Lenguaje de marcado extensible”. En 1996, el Consorcio para el desarrollo de la *World Wide Web* o W3C, sentó las bases para estos desarrollos: se establecieron las ventajas propias del SGML: estructura, extensibilidad y validación, y se creó un grupo de trabajo que estableciera las bases para un nuevo lenguaje de marcado que conservara las ventajas centrales del SGML pero que tuviera la simplicidad del HTML; esto es, un SGML “aerodinámico” para la *web*. El resultado fue que en 1998 se publicaron las especificaciones de la primera versión de XML. En cuanto a dimensiones, la especificación XML resultó ser de menos de

una décima parte de la de SGML, con lo cual puede estimarse el grado de compactación.

Por otra parte, el formato MARC dió origen a otro tipo de metalenguajes con otros propósitos; no con fines de editar en la *Web*, sino de facilitar el registro catalográfico de materiales en la red. Entre ellos el más connotado ha sido sin duda el “Núcleo de Dublín” o *Dublin Core*. El Núcleo remonta sus raíces a la 2ª Conferencia internacional de la *World Wide Web* en Chicago, en octubre de 1994. Yuri Rubinsky de *Soft-Quad*, quién presidía páneles en cuanto al futuro de HTML e instrumentos de autoridad en la *web*, junto con Stuart Weibel y Eric Miller de OCLC, quienes presentaban ponencias acerca de las publicaciones académicas sobre la *Web* así como reflexiones acerca de la entrega de servicios bibliotecarios basados en la red, tuvieron una conversación informal con Terry Noreault, entonces Director de la Oficina de Investigación de OCLC y Joseph Harden, Director del Centro Nacional para Supercómputo (NCSA). Esta discusión sobre la semántica y la *Web* giró alrededor de la dificultad para encontrar recursos en la red, aun entonces con sólo aproximadamente 500,000 objetos direccionables sobre la *Web*; ello debido a la falta de homologación para registrar ese material.

Su reunión inicial condujo a NCSA y OCLC a crear un taller conjunto para discutir sobre la semántica de los metadatos en Dublín, Ohio, en marzo de 1995. En ese evento, llamado simplemente el *OCLC/NCSA Metadata Workshop*, más de 50 expertos discutieron cómo un núcleo semántico de metadatos podría ser útil para clasificar documentos sobre la *web* y facilitar su búsqueda y recuperación. Ellos llamaron al resultado “Núcleo de Metadatos de Dublín” basándose en las conclusiones del taller. Posteriormente, las discusiones y conclusiones se extendieron a múltiples países del mundo y empezaron a ser utilizadas a nivel de recomendación, ganaron mucha aceptación por su simplicidad y ayuda práctica, y se volvieron un estándar *de facto*. En 2001, el formato fue ampliado para incluir tutoriales y ponencias arbitradas y “carteles” para conferencias, y le ofrecieron a la comunidad interesada en los metadatos una mayor oportunidad para aprender e intercambiar ideas, así como promover el desarrollo de estándares de metadatos. El “Núcleo de Dublín” sigue atrayendo a la comunidad mundial, y afilia a más de 1500 organizaciones de más de 50 países con grupos de trabajo

enfocados a múltiples temas relacionados. En 2003 se convirtió en un estándar ISO con el número ISO 15836:2003(E)

MARC y *Dublin Core* están íntimamente relacionados con un conjunto de estándares: “Z39.50” es el estándar nacional norteamericano que define un protocolo para recuperar información de computadora a computadora. El Z39.50 hace posible que un usuario de un sistema busque y recupere información de otros sistemas de computadora que hayan adoptado también registros basados en Z39.50, aun sin saber la sintaxis de búsqueda que es usada por aquellos otros sistemas. El Z39.50 fue aprobado por la Organización Internacional de Estándares de Información (NISO) en 1988 y dio origen al estándar internacional ISO-23950.

4.3.2 Metalenguajes

Si se observa la literatura al respecto, puede notarse que HTML, SGML, XML, Dublin Core y otros entes del manejo de información documental son denominados “lenguajes”, pero, más propiamente dicho, se trata de “metalenguajes”, ¿qué significa realmente este término? es importante aclararlo para poder comprender el contexto de esas herramientas.

El concepto de “metalenguaje” no es simple y pretende establecer las reglas por las cuales un lenguaje dado puede ser definido o examinado *formalmente*; En esencia, se trata de un lenguaje acerca de un lenguaje. Pero existen muchas clase de lenguajes: los lenguajes que los humanos usamos para comunicarnos cotidianamente —español, inglés, francés, etcétera—; los lenguajes para programar una computadora —Fortran, Algol, “C”, Pascal, Cobol, Java, etcétera—; el álgebra, etcétera. Así pues existen muchos metalenguajes; más de los lenguajes que pueden ser creados. Cualquier lenguaje que hablemos las personas incluye varios “lenguajes” o “argots” propios de una profesión o clase social. Distinguimos, por ejemplo, el lenguaje propio de los médicos dentro de un lenguaje dado, o el de las personas de cómputo, etcétera. Los enfoques pueden ser entonces filológicos, matemáticos, computacionales, etcétera. Para fines de este trabajo nos centraremos en los enfoques *documentales*, que son los de nuestro interés.

De esta forma podemos establecer un metalenguaje para definir un lenguaje dado y por tanto, *definiremos entonces de una forma simple al metalenguaje como un medio para definir **formalmente** un lenguaje dado*. Si bien entran también aquí consideraciones de sintaxis, contexto, objetos, etcétera, concentrándonos en el tema que nos ocupa, es decir los documentos, resulta que los metalenguajes —MARC, SGML, HTML, XML— son un medio para describir un lenguaje de marcado; dicho de otra forma, son metalenguajes para definir un lenguaje de marcado de documentos.

Las reglas para definir el marcado de documentos, su sintaxis, sus limitaciones y características están dentro de cada uno de esos entes —MARC, SGML, HTML, XML— es decir, ellos internamente van conformando todo un lenguaje para marcar o codificar documentos. El lenguaje está por lo tanto intrínsecamente contenido en cada uno de ellos, pero la manera de presentar el conjunto de las reglas y definiciones de ese lenguaje es un metalenguaje. Por eso decimos que estas herramientas son metalenguajes, más propiamente dicho, que lenguajes.

Históricamente, la palabra “marcado” —*markup*— tiene su origen en el conjunto de instrucciones que se le daban a un impresor o tipógrafo acerca de cómo debería imprimirse un cierto pasaje de un texto: las negritas, los subrayados, la tipografía, el tipo y tamaño de letra, los símbolos especiales, etcétera, se señalaban con anotaciones o “marcas” al margen del texto, de ahí el nombre. Conforme los textos se fueron automatizando, el término se fue extendiendo hasta cubrir toda clase de códigos de marcado insertados en esos textos electrónicos con objeto de incrustar una serie de características propias de cada sección o parte del texto.

De este concepto se desprende que el marcado es una codificación agregada al texto, gracias a la cual cada porción del mismo se va haciendo explícita y toma forma en un contexto. De hecho, y hablando formalmente, todos los textos han tenido desde hace largo tiempo una manera de hacerse explícitos y entendibles para el lector. Los signos de puntuación, los acentos, los espacios entre palabras, el uso de mayúsculas, los párrafos, los capítulos, los incisos, etcétera, se han agregado a los textos desde hace mucho tiempo para que sus lectores puedan

leer explícitamente un conjunto de letras y les den un sentido. Imagínese el lector un texto en el cual todas las letras están contiguas, sin espacios ni puntuaciones ni nada adicional, y podrá imaginarse la enorme ayuda que estos elementos proporcionan para hacer explícitas estas palabras. Si el lector no se lo imagina, analice entonces este texto:

amamaromaleavivaelamorapapayapaparomaleavivaelamoramama

Como puede observarse, el texto anterior difícilmente puede ser interpretado a golpe de vista; es necesario leerlo y releerlo varias veces para que el cerebro humano, tan avezado en este tipo de actividades, pueda establecer patrones, cortes, analogías, etcétera, y descifre su significado y su contexto. Puesto de esta otra forma:

palíndroma:

*A mamá, Roma le aviva el amor a papá y a papá,
Roma le aviva el amor a mamá*

Es mucho más fácil de interpretar. Ciertas convenciones de espacios y puntuación lo han hecho explícito. Incluso y gracias al título, todos los lectores se percataron que el texto es un palíndroma, es decir, algo que puede ser leído igualmente de izquierda a derecha que de derecha a izquierda. Algunos no se habrán percatado de ello hasta que el título, es decir, cierto *marcado*, lo hizo explícito.

Del mismo modo, y en ello radica su importancia, un lenguaje de marcado debe permitir hacer explícito un texto para los sistemas y las máquinas que los operan electrónicamente, además de decirnos dónde empieza y termina cada palabra, lo cual puede lograrse con espacios, nos indica dónde empieza una parte de un texto en un documento y dónde acaba, qué partes tiene ese documento y cómo se llaman, etcétera. El lenguaje de marcado es entonces un conjunto de convenciones preestablecidas que se usan de manera conjunta y ordenada para marcar o codificar un texto de modo que éste pueda ser entendido explícitamente por máquinas para que éstas, a su vez, se lo puedan hacer explícito a las personas.

Un lenguaje de marcado debe especificar entonces cuáles marcas son permitidas, cuáles son obligatorias u opcionales, cómo deben diferenciarse esas marcas del texto propiamente dicho y, por supuesto, qué significa cada marca y el texto con ella marcado.

Trataremos de ilustrar con un ejemplo los temas anteriormente expuestos. Queremos describir una parte de un documento que es la dirección postal de una persona dada, la cual, como es sabido, está formada a su vez por varios elementos. ¿Cómo describir esta entidad en lenguaje llano pero a la vez *formalmente*? tal vez sería algo como esto:

“Una dirección postal consiste de: una *parte-con-el-nombre*, seguida de una *parte-con-la-dirección*, seguida de una *parte-con-el-código-postal*.

La *parte-con-el-nombre* consiste de: un *nombre-de-pila*, seguido de un *apellido-paterno*, seguido de un *apellido-materno*, si se tiene registrado.

La *parte-con-la-dirección* consiste de: un *nombre-de-calle*, seguido de un *número-exterior*, seguido —si lo hubiese— de un *número-interior*, seguido de un *nombre-de-colonia*,

La *parte-del-código-postal* consiste de: un *nombre-de-ciudad*, seguido de un *nombre-de-entidad-federativa*, seguido de un *código-postal*.

Un *código-postal* consiste de un número de cinco dígitos”.

Puede usarse un metalenguaje para describir esto. Por simplicidad, usamos aquí la “forma de Backus-Naur” —*Backus-Naur Form* o BNF—, un metalenguaje muy utilizado para definir lenguajes de programación desde 1959 y que toma su nombre de sus creadores John Backus y Peter Naur. [Backus-Naur Form, 2000] y [Naur, 1963].

Las bases del metalenguaje son las siguientes; primero los símbolos:

`:=` significa “se define como”

`|` significa “o”

`<>` corchetes angulares usados para encerrar nombres de categorías o “entes”.

{ } corchetes de llave usados para encerrar nombres de categorías opcionales.

“” Texto entre comillas que se integra literalmente tal como está escrito.

<EOL> fin de la línea de definición

La manera de usar la forma de Backus-Naur es la siguiente: se pone del lado izquierdo la entidad a definir y del lado derecho la definición, separadas por el signo de “se define como”. Se trata de pasar de lo general a lo particular y se evitará a toda costa la tautología; es decir que la parte a definir sea igual que la definida. No obstante, *una parte* de lo que está a la derecha como definida puede ser parte de lo que se define. Por ejemplo, en el álgebra, la expresión $n = n+1$ no puede existir, ya que de acuerdo con sus reglas, no hay ningún número que sea igual a él mismo más uno. En BNF $n := n+1$ sí puede existir y significa que n se define como él mismo incrementado en una unidad. Perfecta manera de establecer un contador. Si se observa el ejemplo anterior, se ve que “ n ” forma parte de lo que se define y de lo definido, pero es correcto bajo estas reglas y no es una tautología.

Si en algún momento una definición puede ser definida a su vez en componentes más simples, esa nueva definición se escribe a continuación. De acuerdo con lo anterior, puede definirse más formalmente una dirección postal según la notación de Backus-Naur como:

<dirección-postal> := <parte-con-el-nombre> <parte-con-la-dirección> <parte-con-el-código-postal>

<parte-con-el-nombre> := <nombre-de-pila> <apellido-paterno> {apellido-materno}

<nombre-de-pila> := <nombre> | <inicial> “.” <EOL>
| <nombre> <nombre-de-pila>

<parte-con-la-dirección> := <nombre-de-calle> <número-exterior> {número-interior} <nombre-de-colonia>

`<parte-con-el-código-postal> := <nombre-de-ciudad> <nombre-de-entidad-federativa> <código-postal>`

`<código-postal>:=<dígito><dígito><dígito><dígito><dígito>
<dígito> := 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9`

De acuerdo con este ejemplo, y utilizando la notación BNF puede observarse lo siguiente:

- Se ha definido dirección-postal como un categoría formada por tres partes —nombre, dirección y código— las que deben existir siempre y en ese orden.
- Se ha definido a su vez al nombre como una categoría formada por el nombre-de-pila, seguido por el apellido paterno y, opcionalmente, el apellido materno, y en ese orden.
- Se ha definido también al nombre de pila como una categoría que puede estar formada por dos posibilidades: una de ellas es un nombre o una inicial más un punto obligatorio; la otra posibilidad es compuesta, y puede estar formada por un nombre seguido de lo que habíamos ya utilizado como nombre de pila; es decir, el concepto se vuelve recursivo. Esto es, puedo usar un nombre más otro nombre, o un nombre más una inicial, o una inicial más un nombre, etcétera. Dicho de manera práctica, un nombre de pila puede ser “José” ó “J.” ó “José Luis” ó “J. Luis” ó “José L.” ó “J. L.” Al usar el concepto de recursividad una entidad se define como un nombre seguido de lo que se había utilizado como nombre con anterioridad; es decir, se pueden ir acumulando. Todos los nombres mencionados en el ejemplo caben dentro de esta definición. El símbolo `<EOL>` indica que lo del primer renglón es aparte de lo del segundo y no deben mezclarse.
- Se ha definido también que la dirección está formada por el nombre de una calle, seguido de un número exterior y, opcionalmente, un número interior, seguido de un nombre de colonia, en ese orden.

- Se ha definido que el código postal está formado por el nombre de una ciudad, seguido por el nombre de una entidad federativa y el número de un código postal, en ese orden.
- Se ha definido que un código postal está formado por cinco dígitos consecutivos.
- Se ha definido a un dígito como cualquier guarismo del cero al 9.

Si bien no hemos definido una serie de validaciones y restricciones con respecto a estos posibles datos, ya que se trata de un ejemplo muy sencillo, puede observarse que la definición de un ente llamado dirección postal es posible, precisa y poco ambigua cuando usamos un metalenguaje; en este caso la notación BNF. Esto es más fácil que tratar de explicarlo en lenguaje corriente, ya que habría que estar dando muchas explicaciones sobre cada elemento y además el resultado es una definición que tiene mayor formalidad. Además estas definiciones pueden ser interpretadas por un programa de computadora. Puede hacerse entonces que ese programa escriba las categorías en sitios precisos dentro de una pantalla, hoja o etiqueta de papel: primero el nombre, luego la dirección y luego el código. Puede hacerse un listado sólo de los nombres de las personas y omitirse lo demás; puede hacerse un listado que ordene los nombres alfabéticamente por apellidos, etcétera. Todas estas acciones serían muy difíciles si el nombre y la dirección de una persona fuesen una sola entidad, ante lo cual una máquina se vería imposibilitada de distinguir cada una de esas partes.

Cabe hacer notar que esta notación de Backus-Naur es sólo un ejemplo de un metalenguaje para ilustrar conceptos básicos sobre este tópico. No se usa para describir un texto en la práctica ya que carece de un buen número de elementos necesarios para realizar el encodificado de ellos mismos. Para ello se han creado metalenguajes propios para ese efecto.

4.3.3 HTML —Hyper-Text Markup Language—

Como ya se ha mencionado, a principios de los noventa se creó la *World Wide Web*, *www*, o simplemente *Web*, como un mecanismo para editar y acceder a información a nivel mundial con base en las tecnologías de Internet o red global de computadoras. En un inicio la información que viajaba en la Internet era sólo de tipo texto, y las instrucciones debían ser tecleadas a modo de comandos. El advenimiento de la *Web* le agrega interfaces gráficas a la distribución de información; es decir que una página de la *web* contiene ahora además de textos, imágenes —fotografías, dibujos, diagramas, etcétera—; así como otros elementos de acceso más cómodos para el usuario: botones para seleccionar opciones, menús, etcétera. Agrega también un elemento muy importante: el hipervínculo; es decir una *liga* con otro texto, que puede estar en esa misma página o en cualquier otra, incluso en otra computadora, sistema o hasta otro país. Gracias al hipervínculo, el usuario puede, al posicionarse sobre un punto en la pantalla, dirigirse hacia otro conjunto de información de su interés sobre la *Web*. Posteriormente se agregaron al medio sonidos e imágenes en movimiento, como animaciones o video, para hacer de la edición en la red un ambiente realmente multimedios.

Precisamente para poder efectuar esta edición de información por medio de la *Web* se desarrolla el metalenguaje HTML, que como ya se ha visto nace como un subconjunto o aplicación muy específica de SGML, del cual sólo se deseaban algunas de sus características aquellas que permitieran describir las partes de una página pensando sobre todo en cómo iban a verse sobre la pantalla. Lo principal era el *lay out*, o distribución física de la información sobre la pantalla: dónde deben quedar los encabezados o letreros, su tipo de letra, tamaño y colores; el diseño y colores de los fondos; ubicación, color, tamaño, etcétera, de párrafos sobre la pantalla; la ubicación, tamaño, etcétera, de las imágenes sobre la pantalla; los botones para seleccionar alguna opción de entre un menú; las ligas con otras páginas, etcétera.

Como puede verse, HTML podía definir un cierto número de entidades en una página a la cual podía arribar un usuario, y brindarle acceso a la información. Pero adolece de un serio defecto: los textos en su

interior son manejados solamente como eso, como “textos”; es decir, como un conjunto de palabras contiguas, sin ningún discernimiento de su contexto, su ponderación, sus relaciones, etcétera. Por esta razón al multiplicarse el número de páginas en la red se hizo necesario un mecanismo que nos permitiera saber en dónde podría aparecer algo sobre un tópico de nuestro interés. Nacen así los “buscadores”, con el fin de poder buscar y encontrar algo en la red. El problema es que los textos sólo están marcados como textos, cadenas de palabras sin ninguna ponderación sobre su contexto o sus partes. Las primeras búsquedas sólo hacen un barrido basándose en la *fuerza bruta* de la computadora que trata de identificar las palabras contenidas dentro de esos textos. Todos los que hemos navegado y utilizado la red sabemos el resultado de estos procesos: docenas, cientos y hasta miles de referencias inútiles llenas de “ruido” e irrelevantes para nuestro propósito. Quienes conocen algo de catalogación bibliográfica saben que es muy diferente buscar por ejemplo, a Octavio Paz como autor, como título o como tema en una obra. Peor aún si sólo buscamos por “paz” fuera de todo contexto, pues encontraremos a otras personas con nombre o apellido “Paz”, ciudades, noticias sobre la paz o la falta de ella, documentos, tratados, estudios, etcétera, que nos hablan de este estado social y que nada tienen que ver con nuestro Premio Nóbel, y hasta menciones sobre la paz de los sepulcros.

Peor aún fue cuando las páginas comenzaron a ser puerta de entrada para enormes repositorios de información, digamos un catálogo de biblioteca ¿cómo encontrar la información que está dentro de esa base de datos, pero que físicamente no está en la página *Web*? las palabras para recuperar documentos usadas como llaves están asociadas a la página en sí misma, pero no a los bancos de datos ligados a ellas. Esa información está perdida para la recuperación.

- Los buscadores tuvieron que irse sofisticando para poder entender con este problema y hoy varios mecanismos se han introducido con este fin, como los “metadatos”, los “buscadores inteligentes” y algunos otros aditamentos como: frecuencia de la palabra en una página, proximidad o contigüidad entre ellas, etcétera. Bastante se ha avanzado en este sentido, y debemos

reconocer que en la actualidad hay páginas con mejores índices y buscadores realmente más adecuados para encontrar información pertinente. Pero el problema viene de origen, desde la manera en que la información es preparada y descrita en HTML. Este formato tiene algunas otras deficiencias, entre las cuales las más relevantes son que HTML:

- No es extensible; es decir, el conjunto de etiquetas es cerrado o finito y nadie puede definir sus propias etiquetas para requerimientos específicos. Por ejemplo, en el mundo bibliográfico querríamos tener etiquetas tales como <autor>, <título> o <ISBN> en lugar de la etiqueta genérica <p> de párrafo.
- No permite representar las especificaciones de las estructuras de los datos, como se requiere en la creación y el uso de bases de datos.
- No provee soporte para validar los datos.

HTML, por lo tanto, no maneja los aspectos de contenido y no puede ser utilizado, por ejemplo, para calificar el contenido de páginas Web. HTML sólo se enfoca a la presentación de la información. Más allá de esto cualquier cosa requiere de un considerable esfuerzo de programación adicional, ya sea en forma de añadidos llamados *applets* o de programas.

No obstante lo anterior, sería un grave error menospreciar el HTML. Este estándar ha logrado resultados inusitados en la expansión de la divulgación documental electrónica en los últimos años. Gracias a él la Web ha llegado a ser lo que es hoy en día. HTML reforzó el hecho de que la plataforma de cómputo fuese irrelevante para el intercambio de información, y creó las bases del mecanismo de transporte para mover documentos a lo largo de distintas redes —lo que llegó a ser el protocolo HTTP, *Hyper-Text Transport Protocole*—, y también el esquema de direccionamiento de documentos tanto locales como remotos —lo que llegaría a ser el direccionamiento URL *Universal Resource Locator*—. Bastante se hizo con HTML a pesar de su esbeltez y sus limitaciones; simplemente resulta que las necesidades de manejo documental siguen creciendo y empieza a verse limitado para contender con ellas. Como muchas otras grandes herramientas, el HTML está siendo

rebasado por el mismo dinamismo de la información que contiene la red. El consorcio W3C ha considerado otras opciones que puedan contender con esas demandas y es por ello que SGML es vuelto a tomar en cuenta y nacen nuevos formatos como XML ó XHTML. Sin embargo, por un buen tiempo seguramente se seguirán viendo muchas aplicaciones desarrolladas en HTML. La aparición de XML y su uso seguramente se implantarán por medio de un cambio gradual de los desarrollos actuales y no mediante una sustitución inmediata.

Si bien los esfuerzos del grupo de trabajo del W3C se concentraron durante un tiempo en la definición del XML, como lo hemos ya mencionado, recientemente se liberaron las especificaciones de una primera versión de un metalenguaje denominado XHTML —eXtensible Hyper-Text Markup Language— compatible con XML, cuyas especificaciones han sido liberadas con anterioridad por el mismo consorcio [XHTML 1.0]. En lo personal no me queda muy claro por qué existen dos sucesores de HTML provenientes de la misma fuente. Como ya se ha explicado, HTML se veía cercado por una serie de limitaciones que eran resueltas con base en aplicaciones adicionales, y por ello un grupo de trabajo desarrolla un nuevo metalenguaje que, conteniendo al anterior —HTML—, pueda contender con las necesidades actuales, y por esto se desarrolla y libera así la primera versión de XML; hasta ahí todo suena muy lógico. El por qué el mismo consorcio libera después una nueva versión de HTML y la llama XHTML y es compatible con XML no está claro, y crea mucho mayor confusión en los desarrolladores al no saber éstos si sus nuevas aplicaciones provenientes de HTML deberán hacerse en XHTML o en XML. Pienso que el W3C quiso condescender con diversas facciones de grupos de desarrolladores y aceptó esta nueva versión del HTML convertida a XML con objeto de darle gusto a todos, pero me parece que la versión con más futuro será a la larga la denominada XML.

4.3.4 SGML —Standard Generalized Markup Language—

Como ya se ha comentado, el SGML sirvió de base al HTML. El primero, mucho más completo que su sucesor, además de los aspectos propios

de la presentación del documento, propuso el concepto de crear una o más definiciones de tipos de documentos —DTD, o *Document Type Definition*—, las cuales son descripciones formales de la sintaxis de los documentos que se le asignan a cada tipo de ellos que se desea definir. Por lo tanto, se requiere de una definición DTD para poder interpretar y verificar un documento SGML.

SGML tiene entre sus principales características ventajas; primero, el ser un estándar no propietario; es decir, que no está atado a una marca, patente o compañía en particular, y que es apoyado por un gran número de proveedores de *paquetería* de cómputo. Por ello, un documento que cumpla con los estándares SGML tendrá una vida más larga que uno basado en un estándar propietario. En segundo lugar, los documentos codificados bajo SGML aún son bastante legibles para las personas y por tanto a la vez legibles para los programas de computadora. En tercer lugar, los documentos bajo SGML describen la estructura de los datos y su semántica, y no sólo la manera en que van a ser presentados en la pantalla.

Entre sus desventajas podemos mencionar: primero, el hecho de que es muy generalizado, que incluye especificaciones particulares para todo tipo de documentos y que se vuelve muy complejo; sus especificaciones se extienden por más de 500 páginas. En segundo lugar, muchas de esas especificaciones son irrelevantes para el uso del documento en la *Web*, lo cual las hace superfluas cuando ese es el uso que se pretende dar al documento. Dado que hay muchas opciones dentro de éste, la interoperabilidad entre diversas empresas se reduce sensiblemente.

Por lo anterior, los principales usos de SGML se han dado en ambientes “cerrados” donde es más o menos fácil controlar la estandarización de documentos, tales como los sectores militares, de inteligencia, de manufactura de aeronaves, empresas de publicación, grandes sistemas de archivo, etcétera [SGML ISO-8879, 1986].

4.3.5 XML —eXtensible Markup Language—

En los anteriores capítulos hemos visto ya que las ventajas y desventajas de HTML y SGML se contraponen: HTML es práctico y esbelto, pero

ha sido rebasado por las necesidades del momento en la *Web*, y SGML, siendo mucho más completo tiene muchos elementos imprácticos para su uso en la red. Hemos visto también que el W3C, tratando de contener con esta problemática, favoreció el desarrollo de un nuevo estándar que tratase de tomar “lo mejor de ambos mundos” y es así como nace el nuevo estándar XML. Éste pretende combinar las mayores capacidades descriptivas y semánticas del SGML con lo práctico, sencillo y universal del HTML. Podría afirmarse que se trata de *la versión corregida y aumentada de HTML o un SGML aerodinámico para la Web*. [World Wide Web Consortium. XML. 2002].

En efecto, XML presenta una serie de ventajas prácticas que lo hacen ideal para ir sustituyendo al HTML en la *web*, hélas aquí:

- Sus especificaciones vienen contenidas en sólo 26 páginas.
- Los programas analizadores sintácticos de XML, que se construyen en una computadora dada, no requieren del mencionado descriptor del tipo de documento (DTD) para poder separar un documento en sus componentes.
- No permite ninguna desviación de la sintaxis estándar. Esto quiere decir que todos los documentos codificados en XML pueden ser editados, guardados y enviados sin importar el programa analizador de XML del receptor.
- Los documentos XML pueden proveer *hojas de estilo* que le permiten a los navegadores —Netscape, Explorer, Mozilla, etcétera— convertir los documentos para presentarlos en HTML en la pantalla.
- Los analizadores XML pueden analizar incluso documentos HTML bien formados. Esto permite que la transición de documentos ya existentes en HTML en una empresa que esté evolucionando hacia XML pueda hacerse de manera gradual y programada.

Como ya se ha establecido, XML es un metalenguaje que permite establecer un conjunto de reglas para definir una sintaxis específica, la cual será aplicada solamente en un escenario preestablecido. Los documentos que pueden especificarse entonces bajo XML pueden ser de muy diversas naturalezas, por ejemplo las propias del mundo de las

finanzas, al especificar datos de acciones, cotizaciones, tipos de cambio de divisas, etcétera. Los datos de tipo bibliográfico contenidos en un documento dado; o bien datos de compuestos químicos dentro de una fórmula; o planos y datos sobre la construcción de un avión, etcétera.

Se han presentado una serie de postulados acerca de las principales diferencias entre el estándar HTML y el XML. Tal vez la mejor manera de comprender esto plenamente sea con un pequeño ejemplo que lo ilustre. Tomemos un mínimo trozo del guión dramático *Don Juan Tenorio*, tal como aparece el texto impreso en un libro, sin ningún lenguaje de marcado electrónico:

Don Juan Tenorio
José Zorrilla

PARTE PRIMERA.

ACTO PRIMERO.

LIBERTINAJE Y ESCANDALO.

Hostería de Cristófano Buttarelli. Puerta en el fondo que da a la calle: mesas, jarros y demás utensilios propios de semejante lugar.

ESCENA PRIMERA

Don Juan, con antifaz, sentado a una mesa escribiendo. Buttarelli y Ciutti, a un lado esperando. Al levantarse el telón, se ven pasar por la puerta del fondo máscaras, estudiantes y pueblo con bachones, músicas, etcétera.

DON JUAN:

¡Cuál gritan esos malditos!
Pero ¡mal rayo me parta
si en concluyendo la carta
no pagan caros sus gritos!
(Sigue escribiendo.)

BUTTARELLI:(A Ciutti.)

Buen carnaval.

CIUTTI:(A Buttarelli.)

Buen agosto
para rellenar la arquilla.

BUTTARELLI:¡Quiá! Corre ahora por Sevilla
poco gusto y mucho mosto.
Ni caen aquí buenos peces,
que son casas mal miradas
por gentes acomodadas,
y atropelladas a veces.

Para leer e interpretar este texto de Zorrilla, una persona no requiere texto con marcado. Es capaz de reconocer el significado de cada una de las partes que lo forman basándose en su experiencia y el simple *marcado* tipográfico. Una versión *mu*y simplificada en HTML de este texto se vería así:

<H1> Don Juan Tenorio
José Zorrilla </H1>

<H2>PARTE PRIMERA.
ACTO PRIMERO.
LIBERTINAJE Y ESCANDALO. </H2>

<P>Hostería de Cristóforo Buttarelli. Puerta en el fondo que da a la calle: mesas, jarros y demás utensilios propios de semejante lugar. </P>

<H3>ESCENA PRIMERA</H3>

<P><I>Don Juan, con antifaz, sentado a una mesa escribiendo; Ciutti y Buttarelli a un lado esperando. Al levantarse el telón se ven pasar por la puerta del fondo máscaras, estudiantes y pueblo con hachones, música, etcétera.</I></P>

<P>DON JUAN:

¡Cuál gritan esos malditos!
¡Pero mal rayo me parta
si en concluyendo la carta
no pagan caro sus gritos!
(Sigue escribiendo.)</P>

```
<P><B>BUTTARELLI:</B><I>(A Ciutti:</I>
    Buen carnaval </P>
<P><B>CIUTTI:</B> <I>(A Buttarelli:</I>
    Buen agosto
    para rellenar la arquilla.</P>
<P><B>BUTTARELLI:</B>¡Quiá! Corre ahora por Sevilla
    poco gusto y mucho mosto.
    Ni caen aquí buenos peces,
    que son casas mal miradas
    por gentes acomodadas,
    y atropelladas a veces.
```

La nomenclatura de marcado HTML es la siguiente:

`<>` encierran una etiqueta de marcado.

`</>` encierran el fin de una etiqueta de marcado. Siempre van por pares: `<>` comienza marcado; `</>` termina marcado para esa etiqueta.

“H” significa “Header” o encabezado. Puede haber más de uno y por tanto se numeran: H1 es el “Header 1”. `<H1>` marca de comienzo del encabezado 1. `</H1>` marca del fin del encabezado 1.

“P” significa “Paragraph” o párrafo. `<P>` y `</P>` indican el principio y el fin de un párrafo.

“I” significa “Itálicas”. “B” significa “Bold”, negritas. Puestos en pares para indicar dónde comienza y termina esa característica del texto, con ayuda de la diagonal “/”.

Para facilitar la comprensión de la lectura, sólo se han ilustrado algunas de las etiquetas básicas del marcado HTML. En realidad, una versión un poco más elaborada y real se vería así: —el principio básico es el mismo:

```
<html>
<head>
<title>Don Juan Tenorio</title>
</head>

<body bgcolor="#FFFFFF" text="#000000">
```

```
<table width="90%" border="1" align="center" bordercolor="#CCCCCC">
  <tr>
    <td height="1081">
      <h1 align="center">Don Juan Tenorio</h1>
      <h2 align="center">Jos&eacute; Zorrilla </h2>
      <p><font face="Arial, Helvetica, sans-serif"
        size="3">PARTE PRIMERA. <br>
        ACTO PRIMERO. <br>
        LIBERTINAJE Y ESCANDALO. <br>
        Hoster&iacute;a de Crist&oacute;fano Buttarelli. Puerta en el fondo que
        da a la calle: mesas, jarros y dem&aacute;s utensilios propios de semejan-
        te lugar. <br>
        ESCENA PRIMERA <br>
        Don Juan, con antifaz, sentado a una mesa escribiendo. Buttarelli y
        Ciutti, a un lado esperando. Al levantarse el tel&oacute;n, se ven pasar por
        la puerta del fondo m&aacute;scaras, estudiantes y pueblo con hachones,
        m&uacute;sicas, etc&eacute;tera. <br>
        </font></p>
        <p><font face="Arial, Helvetica, sans-serif"
        size="3"><b>DON JUAN:</b>
        <br> &iexcl;Cu&aacute;l gritan esos malditos!<br>
        Pero &iexcl;mal rayo me parta <br>
        si en concluyendo la carta <br>
        no pagan caros sus gritos! <br>
        (Sigue escribiendo.) <br></font></p>
        <p><font face="Arial, Helvetica, sans-serif"
        size="3"><b>
        BUTTARELLI: </b>(A Ciutti:) <br>
        Buen carnaval. <br>
        </font></p>
        <p><font face="Arial, Helvetica, sans-serif"
        size="3"><b>
        CIUTTI:</b> (A Buttarelli:) <br>
        Buen agosto <br>
        para rellenar la arquilla. <br>
        </font></p>
        <p><font face="Arial, Helvetica, sans-serif"
        size="3"><b>
```

```
BUTTARELLI: </b>&iexcl;Qui&aacute;ute;!
  Corre ahora por Sevilla <br>
  poco gusto y mucho mosto. <br>
  Ni caen aqu&iacute;ute; buenos peces, <br>
  que son casas mal miradas <br>
  por gentes acomodadas, <br>
  y atropelladas a veces. <br>
</font></p>
</td>
</tr>
</table>
</body>
</html>
```

La versión simplificada en XML de este mismo texto podría ser:

```
<TITLE><B> Don Juan Tenorio</B></TITLE>
<AUTHOR><B>José Zorrilla</B><AUTHOR>
```

```
<TITLE>PARTE PRIMERA.
ACTO PRIMERO.
LIBERTINAJE Y ESCANDALO. </TITLE>
```

```
<TITLE>Hostería de Cristóforo Buttarelli. Puerta en el fondo que da
a la calle: mesas, jarros y demás utensilios propios de semejante lugar.
</TITLE>
```

```
<ACT><TITLE>ACTO I</TITLE>
<SCENE><TITLE>ESCENA PRIMERA.
```

```
Don Juan, con antifaz, sentado a una mesa escribiendo; Ciutti y Butta-
relli a un lado esperando. Al levantarse el telón se ven pasar por la puerta
del fondo máscaras, estudiantes y pueblo con hachones, música, etcétera.
</TITLE>
```

```
<SPEECH>
<SPEAKER>DON JUAN:</SPEAKER>
<LINE>
  ¡Cuál gritan esos malditos !
  ¡pero mal rayo me parta
```


Si en concluyendo la carta
No pagan caro sus gritos!
<COMMENT> (Sigue escribiendo.)</COMMENT>
</LINE>
</SPEECH>
<SPEECH>
 <SPEAKER>BUTTARELLI:</SPEAKER>
<LISTENER>(A Ciutti:)</LISTENER>
<LINE>Buen carnaval </LINE>
</SPEECH>
<SPEECH>
 <SPEAKER>CIUTTI:</SPEAKER>
<LISTENER>(A Buttarelli:)</LISTENER>
<LINE>Buen agosto
 para rellenar la arquilla.</LINE>
</SPEECH>
<SPEECH>
 <SPEAKER>BUTTARELLI:</SPEAKER>

<LINE>¡Quiá! Corre ahora por Sevilla
 poco gusto y mucho mosto.
 Ni caen aquí buenos peces,
 que son casas mal miradas
 por gentes acomodadas,
 y atropelladas a veces.
</LINE>
</SPEECH>

Las terminaciones </SCENE> y </ACT> estarán hasta donde se encuentre el final de la escena y el final del acto, por tanto no se escriben aquí. Los marcadores <> y </> se manejan igual que en HTML y por pares.

Si se comparan estos ejemplos marcados se puede notar que en el primer caso, con HTML, los marcados nos indican encabezados, párrafos, itálicas y negritas; es decir, indicaciones para el despliegue del texto, pero nada acerca de las partes o el contexto que van componiendo la obra. En el segundo ejemplo, el de XML, puede observarse que se ha marcado, además del encabezado o título, dónde empieza el acto I;

dónde empieza la escena; dónde empieza y termina el título; dónde comienza y termina cada diálogo —speech—; quién es el actor del diálogo y quién es el que escucha, y cuál es cada línea del diálogo que le pertenece a cada quien. De esta forma, el documento ha quedado marcado con una serie mucho más completa de elementos, la cuál es susceptible de ser analizada también por una computadora.

Supóngase que se desea que la computadora reproduzca por medio de un programa sintetizador de voz los diálogos a través de las bocinas de la misma; por supuesto con voces propias a cada personaje. Con el primer marcado, sería muy difícil lograrlo; con el segundo, la máquina puede, al ir “leyendo” el texto, identificar fácilmente quién es el actor que habla cada línea <SPEAKER>, utilizando siempre la misma voz para ese personaje, y reproducir sólo el texto que se encuentra en <LINE>, ignorando para el parlamento las indicaciones de título, nombre del actor, quién es su interlocutor, etcétera.

A propósito hemos denominado las etiquetas en inglés: *title, act, speech, listener, line*, etcétera. Podríamos haberlas denominado título, acto, diálogo, escucha, línea, etcétera. Sería válido y obviamente funcionaría en nuestro ámbito. Pero ¿qué pasaría si se quisiera que ese diálogo fuera puesto en *Internet* para su uso universal? ¿cómo sabrán las máquinas de otras partes del mundo que <diálogo> es eso, un diálogo? Si bien no hay un estándar establecido en cómo debe llamarse a una etiqueta, es muy probable que en poco tiempo los dramaturgos que quieran poner sus textos en esta forma se inclinen por un nombre universal para la marca, y lo más probable es que sea un nombre en inglés y que por convención, todo el mundo acepte en poco tiempo que “<speech>”, por ejemplo, es siempre un diálogo.

Otro ejemplo de aplicación de este marcado es que una computadora, bajo pedido, puede hacer fácilmente una lista de los personajes de la obra, <SPEAKER> y ponerla en su directorio de palabras para recuperación, con objeto de que un usuario en la red busque y encuentre a tal o cual personaje que aparece en una obra. Hacer esto con el primer marcado sería sumamente difícil.

Cabe hacer notar aquí que en ambos ejemplos no se han utilizado todos los elementos posibles del marcado de cada metalenguaje. Ni con mucho se han agotado todas la etiquetas utilizables. Sólo hemos

seleccionado unas cuantas y simplificado el ejemplo, para ilustrar el concepto. No es el propósito de este documento presentar todos los elementos de marcado de cada formato, porque esto sería inmenso; para ello existen textos a propósito. La idea es resaltar las características descritas a lo largo de este texto.

Resumiendo, podemos afirmar que XML es:

- **Simple** . La especificación completa mide menos de 30 cuartillas. XML ha sido diseñado para facilitar aún más la escritura de programas con respecto a HTML o SGML.
- **Extensible** . Cada quien puede inventar sus propias etiquetas para marcar cualquier tipo de documento, y ser éstas compartidas. De hecho, XML es un metalenguaje que le permite al usuario definir su propio lenguaje de marcado.
- **Un estándar abierto** . XML es SGML. Ello significa que no es necesario saber programar; existen muchas herramientas eficientes que permiten ya sea crearlo, manejarlo o implantarlo en una computadora, y distribuirlo.
- **Eficiente** . XML tiene entes interconstruidos para reutilizar fragmentos de documentos, así, estos sólo tienen que ser transmitidos una vez.
- **Basado en la experiencia** . XML ha sido diseñado por personas que tienen amplia experiencia en los lenguajes de marcado y han capitalizado las enseñanzas que el uso de ellos les ha dado a lo largo de los años.
- **Consensado** . El diseño de XML incluye los puntos de vista de los organismos coordinadores de HTML y SGML, así como los de personas que han desarrollado importantes aplicaciones con estos estándares.
- **Libre**. Nadie tiene la propiedad o patente de XML, ni podrá tenerla, ya que tanto SGML como XML han sido definidos como estándares internacionales. Por lo mismo su uso o desarrollo no implica el pago de ninguna regalía.

- **Inter nacional** .XML tiene interconstruido un soporte para textos en prácticamente todos los alfabetos del mundo, incluyendo técnicas para consignar el lenguaje y/o código del alfabeto utilizado.
- **Listo para ser usado** .Los *browsers* o *navegadores* del *web*, en sus últimas versiones, son capaces de leer especificaciones XML. Los hipervínculos, textos y multimedios pueden ser vistos tal como se hace ahora con un documento HTML.
- **Manejable** .XML incluye métodos para declarar y reforzar las estructuras documentales usadas actualmente, como las de bases de datos.
- **Validable** .XML tiene técnicas que permiten la validación de los documentos involucrados, así que uno puede estar seguro de que los documentos registrados con él son creados correctamente.

Para la última versión del formato, el MARC 21, se incluyó una versión de “MODS” —*Metadata Object Description Schema*— construida por la Oficina de Desarrollo de Red y la Oficina de Estándares MARC de la Biblioteca del Congreso de los EUA, junto con algunos expertos interesados. Éste consiste en un esquema para un conjunto de elementos bibliográficos que puede ser usado para una variedad de objetivos, y en particular para aplicaciones de biblioteca. Como todo esquema XML, el *Esquema de Descripción de Metadatos de Objetos* (MODS) está diseñado para ser capaz de llevar datos seleccionados de registros MARC 21 así como permitir la creación de archivos de descripción de recursos originales. Esto incluye un subconjunto que contiene las etiquetas usuales MARC así como un conjunto de etiquetas basadas en el lenguaje usual de catalogación. Los esquemas “MODS” son expresados usando el metalenguaje de esquemas XML del Consorcio del World Wide Web. En resumen, éste es un esquema ya fabricado por la Biblioteca del Congreso de los EUA [MODS, 2004] y respeta, si uno así desea integrarlo, el formato MARC en nuestros registros. Imaginemos una pequeña ficha simplificada:

PQ6575 D572

Zorrilla, José, 1817-1893. Don Juan Tenorio : Drama Religioso - Fantástico en Dos Partes. Barcelona : Maucci, 1914.
VII, 157 p. 23 cm.
I. Arte Dramático - Teatro - Sevilla. II. España - Siglo XVI

La versión “MODS”; es decir, un registro tipo XML del registro MARC 21 de esta ficha, definido según el esquema de la Biblioteca del Congreso, se vería así:

```
<mods version="3.0" xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-0.xsd">
  <titleInfo>
    <title>Don Juan Tenorio:</title>
    <subTitle>Drama Religioso - Fantástico en Dos Partes</subTitle>
  </titleInfo>
    <name type="personal">
      <namePart>Zorrilla, José</namePart>
      <role>
        <roleTerm type="text">author</roleTerm>
      </role>
    </name>
    <typeOfResource>text</typeOfResource>
    <genre authority="marc">bibliography</genre>
  -
    <originInfo>
  -
    <place>
      <placeTerm authority="marccountry" type="code">nyu</placeTerm>
    </place>
  -
    <place>
      <placeTerm type="text">Barcelona</placeTerm>
    </place>
    <publisher>Maucci</publisher>
    <dateIssued>1914</dateIssued>
```

```
<dateIssued encoding="marc">2004</dateIssued>
<issuance>monographic</issuance>
</originInfo>
-
  <language>
<languageTerm authority="iso639-2b" type="code">spa</language-
Term>
  </language>
-
  <physicalDescription>
<form authority="marcform">print</form>
<extent>vii, 157 p. ; 23 cm.</extent>
  </physicalDescription>
<note type="statement of responsibility">José Zorrilla.</note>
-
  <subject authority="lcsb">
<topic>Arte Dramático</topic>
<topic>Teatro</topic>
<geographic>Sevilla, España</geographic>
</subject>
-
  <subject authority="lcsb">
<geographic>España</geographic>
<topic>Teatro</topic>
<temporal>Siglo XVI.</temporal>
</subject>
-
  <recordInfo>
<recordContentSource>LC</recordContentSource>
<recordCreationDate encoding="marc">990730</recordCreationDate>
<recordChangeDate encoding="iso8601">20000406144503.0
</recordChangeDate>
<recordIdentifier>. PQ6575 D572 </recordIdentifier>
</recordInfo>
</mods>
```

Como puede observarse, éste es un “puente” que se crea entre los típicos registros MARC y los nuevos esquemas XML para enriquecer más la información de los mismos reaprovechando los registros ya creados con anterioridad por MARC, ya que pueden hacerse conversiones automáticas y con autoridad por computadora de registros MARC a registros XML. Nótese el uso en MODS de etiquetas de texto en lugar de numéricas,

una de sus mayores ventajas para la lectura del documento por parte de personas.

Revisemos ahora cual es la relación de estas ventajas al usar XML dentro del contexto de las bibliotecas digitales, en especial las del medio mexicano: sin duda una de las características del establecimiento de colecciones y bibliotecas digitales es que éstas no se limitan sólo a materiales de referencia, sino, cada vez más, a textos completos. Las bibliotecas digitales tratan ahora de establecer los mejores mecanismos para almacenar estos textos completos electrónicos de tal forma que sean totalmente recuperables y explotables por las comunidades académicas, más allá de hacer una simple búsqueda de palabras en el texto y de usar operadores *booleanos*. Ha quedado demostrado que XML tiene un gran número de ventajas nativas para este propósito, y que por ello mismo será el lenguaje de marcado más utilizado en esta década hasta que sea sustituido por algo mejor. No obstante, su misma universalidad y extensibilidad, y las mayores ventajas de XML, introducen una serie de problemas en el entorno en que éste puede ser utilizado.

En efecto, debemos empezar a acotar el problema en nuestro medio bibliotecario; como se ha mencionado XML puede ser utilizado para describir toda clase de documentos: desde los *Rollos del Mar Muerto*, pasando por un expediente médico o un libro, hasta las especificaciones de un avión supersónico. Puede ser usado entonces en los sectores educativo, de la construcción, médico, gubernamental, etcétera. Cada sector puede y debería empezar a sentar las bases para definir sus documentos más relevantes. De hecho, puede existir más de una definición para documentos de un mismo sector o tipo, lo que haría muy difícil su manejo por parte de usuarios distintos a aquellos para los cuales se haya diseñado la información.

Por este motivo en la actualidad se están definiendo descriptores de tipo de documento por grupos sectoriales con intereses afines, de forma que están surgiendo estándares avalados por organismos que garanticen que cualquier usuario que los adopte como suyos, trabaje con las mismas etiquetas e idénticas normatividades, como se hace con el actual HTML. Como ejemplos de esto tenemos CML, *Chemical Markup Language* para el sector químico, MathML, *Mathematical Markup Language* para definir datos matemáticos, SMIL, *Synchronized Multimedia*

Integration Language, para definir presentaciones en recursos multimedia, etcétera. A estos tipos de variantes de documentos se los conoce como *clases* y es un término muy importante.

De hecho, para un mismo tipo de documento pueden existir diversos enfoques y necesidades. Tomemos el caso del documento “libro”. La definición de este documento obedecerá a los intereses particulares del sector interesado. Por ejemplo, para alguien que comercializa libros en la *Web*, su definición de documento sólo incluiría elementos mínimos de identificación: autor, título, año, editorial, ISBN; pero también datos tales como precio, descuento, peso, costo de envío, disponibilidad y/o tiempo de entrega, condiciones de pago, etcétera. Poco que ver con el ambiente de una biblioteca, en donde faltan muchos otros datos con fines de referencia de los documentos, mientras que algunos otros de nuestro ejemplo salen sobrando. Si el documento por definirse es un texto completo, deben agregarse muchos más datos. Las definiciones hechas para un sector serían inútiles para otro.

En el medio académico, y en particular el de las bibliotecas, es necesario entonces comenzar a establecer las definiciones tipo para cada clase de documento, con objeto de satisfacer cabalmente las necesidades de consulta por parte de sus comunidades de usuarios. Cabe recordar que existen además distintos tipos de bibliotecas, y por ende, deberá haber ciertos “matices” en estas definiciones. Estos estándares deberían por supuesto, estar acordes con nuestra realidad mexicana y con las características de nuestra producción editorial, entorno histórico y cultural, idioma, etcétera.

Debemos reflexionar entonces si queremos sentarnos a esperar que estas definiciones sean hechas por otros entornos extranjeros, o por sectores ajenos a la academia y a la biblioteca, o por personal no profesional en el registro documental; o si es el medio bibliotecario académico el que debe comenzar a sentar las bases para establecer estándares en estas definiciones, como ya lo han comenzado a hacer otros sectores en otras partes del mundo. *El sector bibliotecario es el que debería hacer las especificaciones de los documentos para el medio académico*, y sería ideal que fuera el sector bibliotecario mexicano el que lo hiciera para el medio académico mexicano.

Hemos visto ya el despertar y el auge que están tomando las bibliotecas digitales en todo el orbe, y el que comienzan a tener en nuestro país. Empezamos a ver ya algunas colecciones de documentos digitales. Sin embargo, cabe observar que casi todas estas colecciones son presentadas en formato HTML y que casi ninguna de estas colecciones de bibliotecas digitales tiene ya sus definiciones XML, aunque hay que destacar que algunas ya lo están considerando y preparándose para ello. Se venden o distribuyen ya en el medio algunos programas o aplicaciones que se anuncian como “generadores de bibliotecas digitales”, tanto de origen nacional como extranjero. Empiezan a surgir ya las primeras bibliotecas digitales con especificaciones XML. La pregunta crucial es ¿quién crea esas definiciones para los tipos de documentos que manejan esas bibliotecas? Un examen más detallado nos enseñará que por lo general esas definiciones no han sido hechas por personal profesional en registro documental, y por lo tanto cuando mucho representan el punto de vista de un sólo profesional o sector. Las “etiquetas” de marcado muchas veces son sólo réplicas de las establecidas por MARC o AACR2, que han sido buenas para material de referencia pero no para textos completos.

Cabe resaltar aquí que estas deficiencias no se hacen notar de un modo peyorativo alguno. Las personas u organizaciones que están detrás de ello han hecho su mejor esfuerzo con los elementos disponibles hasta este momento y sin ningún estándar preestablecido, y por territorios no explorados. Simplemente deseamos resaltar el hecho y el riesgo que se corre de crear una Babel alrededor del tema, hasta que los mejores estándares desarrollados se vayan filtrando en el medio y la situación se estabilice. Pero ello puede tomar mucho tiempo y costar caro, además de consumir recursos que en nuestro país no sobran y deben ser utilizados racionalmente para el mejor desarrollo de nuestras bibliotecas y colecciones digitales y por ende, de las comunidades académicas a las que éstas sirven.

Es altamente necesario, entonces, comenzar a establecer los estándares mínimos que estas definiciones deben contener para cada tipo de documento orientado hacia el sector académico en general y bibliotecario en particular, considerando no sólo las características de referencia, sino los textos completos, y enfocando todo a nuestro medio

mexicano, a nuestras bibliotecas y a nuestro entorno histórico, social, económico, etcétera. Estas definiciones deben ser efectuadas por personal experto, profesional y multidisciplinario en tareas de registro, recuperación y distribución documental, con el fin de que sean avaladas por las bibliotecas digitales que se están creando y de que sean adoptadas por ellas como suyas. La idea es crear un ambiente de homogeneización y calidad en los registros y las colecciones que se formen.

No se trata de ponerle una camisa de fuerza a las especificaciones de los documentos, sino de crear una definición de calidad sobresaliente como base que pueda ser aceptada con confianza por las bibliotecas digitales de nuestro medio, para que a partir de ella cada biblioteca pueda seguir construyendo sus propias especificaciones.

Capítulo 5

Publicaciones electrónicas y derechos de autor y de propiedad

"Ninguna parte de este documento puede ser copiada a mano, fotografiada, fotocopiada, pintada, dibujada, mecanografiada, capturada, escandida, editada, modificada, reproducida, repetida, almacenada, guardada, memorizada, grabada, esgrafiada, esculpida, mejorada, empeorada, enmendada, aclarada, oscurecida, coloreada, retransmitida, transportada, telepatada, alterada, reformateada, desensamblada, descargada, transferida, compartida, trasvasada, distribuida, complementada, derivada, tratada, combinada, recitada ni platicada, ni en verso ni en prosa, por ningún medio impreso ni virtual, tangible o intangible, ya sea mecánico, físico (newtoniano o cuántico), químico, electrónico, fisicoquímico, telefónico, telegráfico, telefónico, matemático, litográfico, xerográfico, magnético, magnetofónico, óptico, biológico, neurológico, metafísico, psíquico, parapsíquico, telepático, cibernético, biónico, genómico, mnemónico, orgánico, inorgánico, eléctrico, electromecánico, electromagnético, opto-magnético, atómico, hidráulico, neumático, geológico, mágico, teológico, alienígena u otros por inventarse o concebirse por ningún ente basado en carbón o en silicio sin mi expreso consentimiento por escrito como titular de los derechos".

Juan Voutssás. *Mi declaración de copyright llevada a ultranza.*

5.1 DERECHO DE AUTOR Y DERECHO DE COPIA.

Antes de establecer las características de los derechos de propiedad autoral relativos a las publicaciones electrónicas y a las colecciones digitales, conviene establecer algunos puntos básicos de las legislaciones vigentes a nivel mundial en términos de las publicaciones en general, a modo de disponer de un marco de referencia sobre el cual elaborar algunas precisiones.

Históricamente el derecho de autor tiene cerca de dos siglos; no obstante, sus antecedentes se remontan hasta los principios de la era de la imprenta, cuando las primeras normativas aparecen en Venecia a finales del siglo xv en forma de monopolios otorgados por las autoridades a ciertos impresores a cambio de favores y lealtades políticas. Ya como derecho de propiedad se originó en Inglaterra en 1710 con

el *Estatuto de la Reina Ana de Inglaterra*,⁷⁰ y tenía, entre otras cosas, una serie de disposiciones que buscaban por un lado proteger los intereses de los autores asegurándoles su propiedad intelectual y por otro lado, era un mecanismo de control del estado sobre los editores al obligar a los autores a solicitar el registro de sus obras. El periodo original era de catorce años renovable por una vez, y se creaba el “dominio público”. La protección autoral continuó en Francia en 1716; pasó a Dinamarca en 1741 y luego a los Estados Unidos en 1790. Poco a poco todos los países del orbe que tenían actividad editorial fueron adoptando alguna variante de estas disposiciones legales, aunque todos con un enfoque local y con un tinte de control sobre el autor. Francia fue el primer país que instauró una legislación de protección internacional en 1842, la cual se consolidó en 1886 con la *Unión de Berna*, en la que catorce países establecían una legislación de protección internacional. A esta unión se fueron uniendo muchos países paulatinamente hasta desembocar en la Convención de Ginebra para el *copyright* de 1952, la que agrupa a más de 120 países hoy en día. Los Estados Unidos abandonarían un poco su regimen proteccionista local y se adherirían a ella hasta 1988.

México, todavía como Nueva España, utilizó el decreto de 1813 de las Cortes de Cádiz a favor de la propiedad de los autores. La primera Constitución de México de 1824, en su Título III, Sección Quinta del Poder Legislativo, artículo 50, estableció entre las facultades del Congreso: *promover la ilustración, asegurando por tiempo limitado derechos exclusivos a los autores por sus respectivas obras*. En 1846 el Ministerio de Instrucción Pública publicó el *Decreto sobre Propiedad Literaria*; dicho instrumento legal constaba de 18 artículos e integró ya el derecho de autor al derecho de propiedad; el decreto se incorporó el 8 de diciembre de 1870 al Código Civil de ese año. El Código Civil de 1884 introdujo ligeros cambios a lo ya legislado sin añadir nada sustancial ya que al igual que en el Código Civil de 1870,

⁷⁰ *An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Author's or Purchasers of Such Copies*. Queen Anne of Great Britain, c. 19 (1709). Disponible noviembre 3, 2005 en: http://press-pubs.uchicago.edu/founders/documents/a1_8_8s2.html

se considera al Derecho de Autor como un derecho real de propiedad de acuerdo con los criterios del siglo XIX. La subsecuente Constitución de 1917 incorporó el Derecho de Autor en su artículo 28. Posteriormente, el Código Civil de 1928, reguló, en tres capítulos, —*artículos del 1181 al 1280 inclusive*— todo lo concerniente al tema autoral.

México participó con otros veinte países más de América en la *Conferencia Interamericana de Expertos para la Protección de los Derechos de Autor, Unión Panamericana*, celebrada en 1946 en Washington, D.C. En ese evento se firmó la *Convención Interamericana sobre el Derecho de Autor de Obras Literarias, Científicas y Artísticas*. Para hacer congruente el derecho de autor mexicano con los compromisos adquiridos en esa Convención, se expidió el 31 de diciembre de 1947 la primera *Ley Federal del Derecho de Autor*.

El 29 de diciembre de 1956, reflejando la normativa de la *Convención Universal sobre Derechos de Autor* de 1952, se expidió la segunda ley sobre la materia, que trató de corregir errores y llenar lagunas de su antecesora; cabe mencionar que bajo esta ley se crea la Dirección General del Derecho de Autor. Posteriormente, la ley de 1963 recibe ascendiente de las ideas contenidas en el *Convenio de Berna* para la Protección de las Obras Literarias y Artísticas en su reunión de Bruselas de 1948, amén de la normativa contenida en la Convención Internacional sobre la *Protección de los Artistas Intérpretes o Ejecutantes, Productores de Fonogramas y Organismos de Radiodifusión*, comúnmente conocida como *Convención de Roma de 1961*. Finalmente, el 24 de Diciembre de 1996 se expide la ley autoral vigente, la que refleja las tendencias neoliberales consagradas en las negociaciones del GATT y finalmente la normativa impuesta en la Organización Mundial de Comercio (OMC), que ya apuntaba su influencia a través del Tratado de Libre Comercio para América del Norte (TLC). La última reforma a ley mexicana data del 23 de julio del 2003.

En términos generales, el derecho de autor o *copyright* es un conjunto de disposiciones legales que le otorgan al propietario de esos derechos ciertas prerrogativas, como el ser conocido y reconocido por su obra, lucrar con ella, así como hacer y permitirlo a otros lo siguiente:

- Reproducir la obra en alguna forma de copia, y distribuirla por algún mecanismo como la venta, el alquiler, el préstamo, etcétera.
- Preparar trabajos derivados basados en la obra.
- Presentar, mostrar o interpretar la obra públicamente, como en el caso de obras artísticas o literarias.

Conviene establecer una definición general del *copyright*, como es conocido internacionalmente y de los *derechos de autor*, ya que la confusión entre ellos conduce a ciertas imprecisiones. Una definición muy sencilla, la podemos ver en esta traducción del manual de Stephen Fishman,⁷¹ *The Copyright Handbook*.

El ‘copyright’ es un dispositivo legal que provee al creador de una obra artística o literaria, o de un trabajo que compila información o ideas, el derecho de controlar cómo es utilizada esa obra. El propósito del ‘copyright’ es promover el progreso del conocimiento dando al autor de una obra un incentivo económico para crear nuevas obras.

Según David Rangel:⁷²

Bajo el nombre de derecho de autor se designa al conjunto de prerrogativas que las leyes reconocen y confieren a los creadores de obras intelectuales externadas mediante la escritura, la imprenta, la palabra hablada, la música, el dibujo, la pintura, la escultura, el grabado, la fotografía, el cinematógrafo, la radio, la televisión, el disco, el casete, el videocasete y por cualquier medio de comunicación”.

Cabe resaltar que el concepto de derecho de autor como concepto jurídico es diferente del *copyright*. El primero proviene del derecho romano-canónico, típico de nuestros países iberoamericanos, y tiene que ver más con el aspecto subjetivo del acto de creación de una obra por un autor dado; el derecho moral consagrado en estas legislaciones tiene su origen en el humanismo y la revolución francesa, de donde se

71 Fishman, Stephen. 2004. *The Copyright Handbook: How to Protect & Use Written Works*. 8th Edition. ISBN: 1-4133-0091-X. 512 p.

72 Rangel, David. 1998. *Derecho Intelectual*. México: McGraw-Hill, p. 111.

extiende a España y de ahí a los países latinoamericanos y a algunos otros de Europa. Su esencia reside en que la obra forma parte integrante de la personalidad del autor: es una creación de su espíritu; el fruto de su pensamiento, de manera que no puede ser disociado enteramente de aquél, por lo que incluso cuando ha cedido sus derechos patrimoniales sobre la obra ésta continúa asociada con él y en cierta medida bajo su dependencia. En resumen, *es un derecho del ser*. En contraste, el *copyright* proviene del derecho consuetudinario anglosajón —*common law*—. Ahí, los derechos morales no son primordiales y su esencia está más intrínsecamente ligada al derecho de copiar o explotar una obra; es decir, su enfoque es más comercial. De hecho, el derecho moral se introduce en la ley inglesa hasta 1988 y no es perpetuo; en la legislación de los E.U.A. aún no se ha reconocido ni se vislumbra la intención de hacerlo. Ello se debe a que en esos países el derecho de autor es básicamente un derecho de propiedad, de manera que una obra intelectual es en principio un artículo de consumo, algo con lo que se puede comerciar bajo el control de la persona o institución que tiene los derechos sobre él. En resumen, *es un derecho del tener*. Si bien son usados como sinónimos, existen estas diferencias de fondo que, como se ve, conviene dejar muy claras.

J.A.L. Sterling ha resumido las principales diferencias entre los derechos de autor europeos y el *copyright* norteamericano [Sterling, 2000]:

- El enfoque del *copyright* se basa fundamentalmente en consideraciones económicas, mientras que el de derechos de autor está vinculado a un concepto de los derechos de la persona.
- En el enfoque del *copyright*, es posible que el autor sea tanto una persona física como moral, mientras que en el ámbito de los derechos de autor se entiende que se trata siempre de una actividad surgida de una persona física.
- En el enfoque del *copyright* el reconocimiento legal de los derechos morales de un autor ha tenido un desarrollo casi inexistente. En contraste, bajo el enfoque de derechos de autor, los derechos morales ocupan una posición preeminente y existe una tradición de alto nivel de protección de tales derechos.

- La fijación en un soporte material de la obra es generalmente esencial en el enfoque del *copyright*. En cambio, bajo el enfoque de derechos de autor no es indispensable para la protección de la obra.
- Cuando se trata de obras cinematográficas, en el enfoque del *copyright* el propietario inicial del derecho puede ser una persona jurídica, por ejemplo, una compañía productora, pero en el enfoque de derechos de autor los propietarios deben ser las personas que contribuyeron a la creación del filme.
- Con el *copyright* el empresario puede ser el propietario inicial del *copyright*, mientras que con los derechos de autor, la regla general es que el empleado es el propietario inicial del derecho de autor, aunque éste pueda ser cedido, mediante contrato, al empresario.
- No es común bajo el *copyright* que los contratos contengan previsiones pormenorizadas sobre la publicación, mientras que tales regulaciones, muy detalladas, son regla en el enfoque de derechos de autor.
- En cuanto a los derechos conexos a los de autor, en Estados Unidos sólo los trabajos originales de autor se protegen mediante *copyright*. Los derechos conexos gozan de una protección minorada o incluso, a veces, de ninguna en especial. En el enfoque de derechos de autor, en cambio, está generalmente clara la distinción entre los derechos del autor en las obras de creación y los derechos conexos o vecinos de los ejecutantes, productores de fonogramas, productores televisivos u otros.

5.1.1 La Ley Federal del Derecho de Autor en México.

En la “Ley Federal del Derecho de Autor” de México [México. Ley Federal del Derecho de Autor, 1996], en sus artículos 11 y 13, se establecen las diferencias explicadas bajo la siguiente forma:

“El derecho de autor es el reconocimiento que hace el estado en favor de todo creador de obras literarias y artísticas previstas en esta ley, en

virtud del cual otorga su protección para que el autor goce de prerrogativas y privilegios exclusivos de carácter personal y patrimonial. Los primeros integran el llamado Derecho moral y los segundos, el Patrimonial. Las ramas de autoría que están amparadas por esta ley son:

- Literaria;
- Musical, con o sin letra;
- Dramática;
- Danza;
- Pictórica o de dibujo;
- Escultórica y de carácter plástico;
- Caricatura e historieta;
- Arquitectónica;
- Cinematográfica y demás obras audiovisuales;
- Programas de radio y televisión;
- Programas de cómputo;
- Fotográfica;
- Obras de arte aplicado que incluyen el diseño gráfico o textil, y
- De compilación, integrada por las colecciones de obras, tales como las enciclopedias, las antologías, y de obras u otros elementos como las bases de datos, siempre que dichas colecciones, por su selección o la disposición de su contenido o materias, constituyan una creación intelectual.
- Las demás obras que por analogía puedan considerarse obras literarias o artísticas se incluirán en la rama que les sea más afín a su naturaleza”.

Continuando con la ley mexicana, cabe resaltar lo mencionado en sus artículos 18, 19, 20 y 21:

“El autor es el único, primigenio y perpetuo titular de los derechos morales sobre las obras de su creación. El derecho moral se considera unido al autor y es inalienable, imprescriptible, irrenunciable e inembargable. Corresponde al propio creador de la obra y a sus herederos el ejercicio del derecho moral. Los titulares de los derechos morales podrán en todo tiempo:

- Determinar si su obra ha de ser divulgada y en qué forma, o la de mantenerla inédita;
- Exigir el reconocimiento de su calidad de autor respecto de la obra por él creada y la de disponer que su divulgación se efectúe como obra anónima o seudónima.
- Exigir respeto a la obra, oponiéndose a cualquier deformación, mutilación u otra modificación de ella, así como a toda acción o atentado a la misma que cause demérito de ella o perjuicio a la reputación de su autor;
- Modificar su obra;
- Retirar su obra del comercio, y
- Oponerse a que se le atribuya al autor una obra que no es de su creación”.

Cabe subrayar que en la ley mexicana vigente emitida en 1996, como en muchas otras legislaciones Iberoamericanas, se hace ya una clara distinción de los derechos de autor en su conjunto dividiéndolos en derechos morales y patrimoniales. El derecho moral, como se consignó, está unido al autor y es inalienable, imprescriptible, irrenunciable e inembargable. Los derechos patrimoniales tienen que ver con la explotación comercial y éstos por lo mismo son susceptibles de transferirse a otra persona física o moral que considere pertinente; lo cual puede hacer a través de permisos, licencias, convenios o contratos, en los que se establece la totalidad o parcialidad de la cesión así como el lapso de ella, y si es a cambio de un beneficio pecuniario o gratuitamente. Como puede observarse el derecho de autor mexicano actual está formado en realidad por dos derechos: el moral y el patrimonial. El patrimonial es el que se parece mucho al *copyright* americano. Así, de acuerdo con la definición de la ley mexicana actual, el término de *copyright* debería ser traducido como *derecho de copia* y no como *derecho de autor*, como comúnmente hacemos, ya que como hemos establecido, hay diferencias de semántica sobre ellos y esto induce a confusiones: y no son tan sólo de enfoque. Las diferencias son de gran trascendencia.

Las bibliotecas, salvo que se dediquen a la publicación de materiales, tienen que ver más bien con los derechos patrimoniales o de *copyright*

que con los derechos morales de los autores. Conviene entonces profundizar en el estudio de éstos últimos, sobre todo las restricciones al uso de esos derechos patrimoniales. Continuando con la ley autoral mexicana, su artículo 27 establece:

“Los titulares de los derechos patrimoniales podrán autorizar o prohibir:

- La reproducción, publicación, edición o fijación material de una obra en copias o ejemplares efectuada por cualquier medio, ya sea impreso, fonográfico, gráfico, plástico, audiovisual, electrónico fotográfico u otro similar;
- La comunicación pública de su obra a través de cualquiera de las siguientes maneras:
 - A) la representación, recitación y ejecución pública en el caso de las obras literarias y artísticas;
 - B) la exhibición pública por cualquier medio o procedimiento, en el caso de obras literarias y artísticas, y
 - C) el acceso público por medio de la telecomunicación;
- La transmisión pública o radiodifusión de sus obras, en cualquier modalidad, incluyendo la transmisión o retransmisión de las obras por: A) cable; B) fibra óptica; C) microondas; D) vía satélite, o E) cualquier otro medio análogo;
- La distribución de la obra, incluyendo la venta u otras formas de transmisión de la propiedad de los soportes materiales que la contengan, así como cualquier forma de transmisión de uso o explotación, cuando la distribución se lleve a cabo mediante venta. Este derecho de oposición se entenderá agotado efectuada la primera venta, salvo en el caso expresamente contemplado en el artículo 104 de esta ley;
- La importación al territorio nacional de copias de la obra hechas sin su autorización.
- La divulgación de obras derivadas, en cualquiera de sus modalidades, tales como la traducción, adaptación, paráfrasis, arreglos y transformaciones, y

- Cualquier utilización pública de la obra salvo en los casos expresamente establecidos en esta ley.”

Nótese el párrafo IV de este artículo, donde he subrayado que “este derecho de oposición se entenderá agotado efectuada la primera venta”. Esta proposición es de singular importancia aunque debido a su redacción legal pasa casi inadvertida. En realidad y gracias a ella, todos podemos regalar, prestar o vender un libro, una revista, un disco, etcétera, que hemos adquirido y deseamos transferir a otra persona, sin caer en falta ante la ley. Significa que el autor puede oponerse a que su obra se venda o distribuya por primera vez, pero una vez que él ha consentido a ello el derecho de oposición cesa, y el autor no puede impedir que el ejemplar de la obra se distribuya. No obstante, queda entendido que al venderla o regalarla entregamos la copia por la cual ya habíamos pagado y que al transferírsela a otra persona **ya no** tenemos esa copia en nuestro poder, y no podemos hacer más copias ni uso de ella. Este fin del *derecho de oposición* o *doctrina de la primera venta* se vuelve sumamente importante y delicado en las colecciones digitales, como veremos más adelante. La excepción marcada en el artículo 104 se refiere a programas de computadora, en los cuales las más de las veces el pago de uso no nos permite transferírselo a un tercero.

Analicemos ahora las limitaciones a los derechos patrimoniales, aspecto de suma importancia para los bibliotecarios, ya que son precisamente estas limitaciones las que nos permiten distribuir el material a terceros desde las bibliotecas.

La ley mexicana, en su capítulo II, artículo 148 establece a la letra:

Capítulo II. Artículo 148. De la limitación a los derechos patrimoniales

“Las obras literarias y artísticas ya divulgadas podrán utilizarse, siempre que no se afecte la explotación normal de la obra, sin autorización del titular del derecho patrimonial y sin remuneración, citando

invariablemente la fuente y sin alterar la obra, sólo en los siguientes casos:

- Cita de textos, siempre que la cantidad tomada no pueda considerarse como una reproducción simulada y sustancial del contenido de la obra;
- Reproducción de artículos, fotografías, ilustraciones y comentarios referentes a acontecimientos de actualidad, publicados por la prensa o difundidos por la radio o la televisión, o cualquier otro medio de difusión, si esto no hubiere sido expresamente prohibido por el titular del derecho;
- Reproducción de partes de la obra, para la crítica e investigación científica, literaria o artística;
- Reproducción por una sola vez, y en un solo ejemplar, de una obra literaria o artística, para uso personal y privado de quien la hace y sin fines de lucro. Las personas morales no podrán valerse de lo dispuesto en esta fracción salvo que se trate de una institución educativa, de investigación, o que no esté dedicada a actividades mercantiles;
- Reproducción de una sola copia, por parte de un archivo o biblioteca, por razones de seguridad y preservación, y que se encuentre agotada, descatalogada y en peligro de desaparecer;
- Reproducción para constancia en un procedimiento judicial o administrativo, y
- Reproducción, comunicación y distribución por medio de dibujos, pinturas, fotografías y procedimientos audiovisuales de las obras que sean visibles desde lugares públicos.”

Estos siete casos son los únicos que permiten realizar copias o distribuir documentos, además de aquellos documentos que se encuentren en el dominio público, ya sea porque así ha sido establecido **expresamente** desde un principio, o porque la vigencia de los derechos patrimoniales ha expirado.

Conviene establecer entonces, para poder completar este conjunto de posibilidades de copia, cuáles se encuentran en el segundo caso; al efecto se presenta una traducción mía de la tabla de Gasaway al efecto:

FECHA DE LA OBRA	RÉGIMEN DE PROTECCIÓN	TÉRMINO
Publicada antes de 1923	Está en dominio público	ninguno
Publicada entre 1923-1963	Protegida cuando se publicó con aviso de restricción	28 años + pudo ser renovada por 47 años + 20 años de otra renovación; si no se renovó, está en dominio público.
Publicada entre 1964-1977	Protegida cuando se publicó con aviso de restricción	28 años de primer término; renovación automática por otros 67 años.
Creada antes del 1-1-1978 pero no publicada	Protegida desde el 1-1-1978	Vida del autor + 70 años o al 31-12-2002, lo que sea mayor*.
Creada antes del 1-1-1978 pero publicada entre esa fecha y 31-12-2002	Protegida desde el 1-1-1978	Vida del autor + 70 años o 31-12-2047, lo que sea mayor*.
Creada del 1-1-1978 en adelante	Desde que la obra es fijada en medio tangible	Vida del autor + 70 años (si es autor corporativo lo más corto entre 95 años desde su publicación, o 120 años de la creación)*

* Cuando se establece la protección de por vida del autor deberá considerarse la del más longevo si es trabajo colectivo. [Gasaway, 2001]

Sumando estos dos conjuntos sabemos entonces cuál es el material que puede ser copiado y redistribuido sin problema, si bien la lista en cada caso es inmensa. Cabe resaltar que ésta es una tabla que sigue la legislación de los EUA, pero es muy semejante a la de la Unión Europea y ha sido imitada por muchos países; dado que los materiales que se requiere copiar no serán solo mexicanos conviene utilizar normativas de corte internacional y de aceptación general a nivel mundial.

Es importante analizar estos plazos de duración de los derechos patrimoniales en sus versiones más recientes. Hago aquí un resumen de lo expresado por Eduardo de la Parra, catedrático de propiedad intelectual en la Facultad de Derecho de la UNAM [De la Parra, 2004]:

La Convención de Berna establece una recomendación de 50 años después de la muerte del autor para que los documentos implicados pasen a ser de dominio público. Después de muchos debates, la Unión Europea llegó en 1993 a la conclusión de que 70 años después de la vida del autor es el plazo ideal de extensión de estos derechos.

Casualmente, es exactamente el mismo plazo que establecieron poco después los Estados Unidos, a través de la *Sonny Bono Copyright Term Extension Act* del 7 de Octubre de 1998, también conocida como la *Ley Disney* ya que el principal interesado en promoverla fue la Corporación Disney, pues algunos de sus personajes más famosos estaban por entrar al dominio público en el 2004. Con esta extensión envió ese paso hacia el dominio público hasta el año 2024, cuando sin duda obtendrán otra extensión por otros veinte años. Originalmente el periodo de validez de los derechos comerciales de autor en los Estados Unidos comenzó siendo de 14 años, tal como en Inglaterra, pero éste parece ser infinitamente elástico ya que ha ido extendiéndose desde mediados del siglo XX, de forma que los derechos de Mickey Mouse y otros contemporáneos podrían no expirar nunca. El último paso en esta tendencia fue la muy discutida *Ley Disney*, que aumento los derechos para trabajos *por contrato* a 95 años desde su fecha de publicación. Como dato curioso, Carlo Collodi, autor de *Pinocho* murió en 1890; Walt Disney filmó esta película animada en 1940. Lewis Carroll, autor de *Alicia en el País de las Maravillas* murió en 1898; Disney filmó esa película animada en 1951. —Disney no pagó nada por el uso de esas historias, en ese entonces estaban en el dominio público—. Si las leyes actuales de derechos de autor de los Estados Unidos hubiesen estado vigentes en la década de los cuarenta, la Compañía Disney, el gran promotor de las extensiones de derechos de propiedad, hubiese tenido que esperar varias décadas para filmar cada una de esas películas, o haber pagado a los herederos de los autores unas buenas sumas. Al final, el problema y el daño no son que Mickey Mouse siga bajo protección de derechos, sino que debido a esa ley miles y miles de obras que estaban por pasar al dominio público han quedado bloqueadas por otros veinte años.

México por su parte, en su ley vigente del año 1996 estableció derechos patrimoniales por 75 años adicionales a la vida del autor, los cuales ya en ese entonces estaban por encima de los aceptados a nivel general en los EUA y Europa. Absurdamente, con la reforma a la ley autoral del 2003 se extienden los derechos patrimoniales a 100 años más allá de la vida del autor, convirtiéndose México en el país que tiene los derechos más excesivos en este sentido a nivel mundial, por

encima de Costa de Marfil, que ostentaba el *record*, con 99 años. Así, México extiende *generosamente* estos derechos al doble de lo recomendado en la Convención de Berna, y casi 50% más de lo establecido legalmente por la unión americana y la unión europea, mercados atractivos para nuestras publicaciones. Al final el esfuerzo será estéril, ya que la protección en esos ámbitos internacionales seguirá limitada a 70 años, lo cual quiere decir que después de ese plazo cualquiera puede editar y comercializar las obras mexicanas fuera de México sin pagar regalías en todo el mundo y sólo en México no podrán ser reeditadas.

Esto suena absurdo en un país donde el acceso a la cultura es de por sí limitado y elitista. Y significa que durante esos 25 años adicionales, será más caro adquirir una obra mexicana en México que en el extranjero, ya que para editar y comercializar una obra aquí sí habrá que pagar derechos y fuera de México no, así que se estará encareciendo inútilmente el acceso a la cultura en México sin lograr ninguna protección adicional. No existe ninguna explicación lógica al respecto de esta decisión del poder legislativo mexicano ya que su antecedente, la extensión de 50 a 70 años en la unión americana, causó uno de los más feroces debates que se recuerden a nivel legal en ese país durante años, y tuvo que ser decidida finalmente por la Suprema Corte de Justicia de esa nación. Me pregunto si los legisladores mexicanos estaban al tanto de ello ya que aquí ellos ni cuestionaron, debatieron o dudaron del punto en cuestión al extender el lapso de 75 a 100 años. Al parecer, la única razón que los legisladores mexicanos pudieron argumentar es que el llamado “Catálogo de Oro” de la música mexicana estaba *a punto de pasar al dominio público y que el plazo de 75 años resultaba insuficiente para el gran volumen de obras que podrían quedar desprotegidas*. Todo indica que esta reforma a la ley se llevó a cabo por esa razón, que obviamente no beneficia a los autores de la música contenida en ese catálogo, ni a sus descendientes más cercanos, ni al gran público o al patrimonio cultural musical mexicano; sino únicamente a las sociedades de gestión autoral.

La flamante reforma a la ley tampoco indica en sus transitorios qué sucede con las obras que ya hubiesen pasado al dominio público antes de la citada reforma y que volvieron a entrar en el plazo extra de 25 años adicionales: ¿se restablecen los derechos patrimoniales de esas

obras? ¿esas obras continúan en el dominio público y la reforma no es retroactiva a ellas? ¿si alguien ya había hecho uso de alguna de ellas debe dejar de usarla so pena de transgredir la ley? —Posteriormente indagué que la interpretación a esa parte de la ley es que no es retroactiva y que las que salieron, sí pasaron a formar parte del dominio público—.

5.1.2 La relación hombre-libro y el equilibrio entre editores y usuarios

Volviendo al punto del derecho de oposición o doctrina de la primera venta, es necesario aclarar cuál es su alcance en el mundo electrónico, ya que de entrada no es obvio cuál es el derecho que un comprador tiene de transferir una obra por la que ha pagado. Históricamente, tampoco fue una característica obvia en las publicaciones en papel, y por ello había que estipularlo claramente. Su origen se remonta a la primera legislación autoral inglesa de 1710; de ahí pasó a las legislaciones de las colonias anglosajonas americanas en 1790, en donde fue cuestionada a lo largo del tiempo por una serie de autores, pero en la mayoría de los casos fue ratificada por los tribunales. Esta doctrina establece que *“el derecho de oposición del autor termina cuando una copia específica de la obra es vendida”*; ello significa, como ya se mencionó, que una persona que compró una copia de un libro puede hacer con ella lo que desee: regalársela a un individuo o institución, prestarla, tirarla, venderla, subastarla, etcétera, con la única limitante de no copiarla; de hecho, ésta es la esencia misma de cómo compartimos información en este mundo.

Esta doctrina es la que les ha permitido también a las bibliotecas prestar libros. Si ellas tuvieran que pagar una regalía cada vez que le prestan un libro a un usuario seguramente tendrían que repercutírsela a éste, y las bibliotecas y la educación no hubieran sido lo que son hoy. Cabe hacer notar que en algunas épocas, en algunos países, sí llegó a haber una “regalía” por el préstamo de un libro. Afortunadamente, se establecieron también muchas bibliotecas en países nacientes en el siglo XVIII y XIX con el concepto de “gratuidad”, el cual permeó a prácticamente

todas a nivel mundial. Si hoy tuviéramos el modelo del pago por el préstamo, la única y mínima diferencia entre una librería y una biblioteca sería que una vende y la otra renta los libros. No obstante, las bibliotecas han pagado caro a los editores por ese servicio para sus usuarios. Desde hace varias décadas, las suscripciones a revistas han sido visiblemente más caras para bibliotecas que para personas, y en algunos casos esto sucede también con otras publicaciones.

Desde 1909, la Suprema Corte de Justicia de los EUA ratificó, y sigue incluyendo, hoy en día la premisa constitucional en las leyes de derecho de autor acerca de que *el propósito del copyright es promover la ciencia y las artes útiles a través de la difusión del conocimiento y que el autor se beneficia de la distribución ya que ello disemina su obra, creando además de nuevo conocimiento, nuevas ventas*. Todas las distintas actas de la ley del derecho de autor americano la han incluido, también la actual, y ha sido recogida por múltiples legislaciones de muchos países. En el caso de la ley mexicana también podemos observarla, como ya hemos mencionado.

No obstante y como ha sucedido a lo largo de la historia, de tiempo en tiempo los editores arremeten contra ella. En 1998 lograron una nueva extensión de los derechos del *copyright* por otros veinte años en los Estados Unidos, y en 2003 por otros veinticinco en México; de hecho, cada veinte años logran una nueva extensión por otros veinte años. Nuevamente el asunto ha salido a colación con recientes tipos de publicaciones, sobre todo las electrónicas. En la práctica, en paquetes de computadora y en videos la doctrina de la primera venta se ha reducido enormemente. De aquí surgen las preguntas ¿cuál es el derecho de primera venta en un documento electrónico? ¿cómo se aplica esto en una biblioteca digital? Las respuestas no son fáciles. La capacidad de reproducción y distribución masiva de la *Internet* es una característica que hay que tener en cuenta. Sin duda crea temor, no sin fundamento, entre autores y editores este asunto acerca del control y las regalías de sus obras. Para las bibliotecas de todo el mundo, este principio continúa siendo crítico para poder realizar su tarea dentro de gastos razonables. Obviamente es del interés de estas instituciones que esta doctrina continúe para este tipo de publicaciones.

El “Acta del Derecho de Copia Digital del Milenio” de 1998 —*Digital Millennium Copyright Act o DMCA*— fue una instancia que representó un esfuerzo del Congreso de los EUA para poner en práctica obligaciones de tratados de ese país y promover las leyes de propiedad intelectual internacionalmente en la era digital. Pero como el Congreso reconoció, *la única cosa que permanece constante es el cambio*. La promulgación del DMCA era sólo el principio de una evaluación en curso que hacía el Congreso sobre la relación entre el cambio tecnológico y la ley de propiedad intelectual estadounidense. Este Informe del Registro de *Copyrights* ha sido recomendado como base para futuros trabajos en el Congreso y rige como guía desde entonces. [United States of America, House of Congress, 1998].

Uno de los elementos destacables en este extenso reporte dice:

“nuestro mandato debía evaluar los efectos de las enmiendas hechas por el título I de la DMCA y el desarrollo del comercio electrónico y las tecnologías asociadas y emergentes sobre la operación de las secciones relativas al título [...] expresamente, este informe se concentra en tres propuestas que nos fueron presentadas durante nuestras consultas: la creación de una **‘doctrina de la primera venta digital’**; la creación de una excepción que **permita crear ciertas copias incidentales temporalmente** y la ampliación de excepciones a el archivado de copias de programas de cómputo”. —Los subrayados en negritas son míos y los resalto porque los utilizaré más adelante—.

Desgraciadamente parece que las buenas intenciones del *Acta del Milenio* se quedaron en eso, en buenas intenciones. Debido a su negligente o malintencionada redacción actual esta disposición legal es actualmente muy cuestionada y ha sido motivo de múltiples debates en varios foros ya que se la acusa —y no sin razón— de ser una normatividad sobreprotectora de los derechos de los productores de material pero que limita y lesiona de manera sensible los derechos de los consumidores y usuarios de información. Por lo mismo fue impugnada en tribunales y hasta se debatió su legalidad en la Suprema Corte de Justicia de ese país; logró pasar sólo después de largos y feroces debates.

Javier Villate resume nítidamente lo anterior cuando afirma:

“...Hoy en día, los autores no son los principales beneficiarios de la propiedad intelectual, salvo unos pocos casos que confirman la regla. La mayoría de los autores, artistas, escritores y demás son explotados por la industria de los contenidos. La propiedad intelectual ya no cumple aquella misión de incentivar la creación y es una broma considerarla como un derecho de la persona. En la práctica, los intereses económicos de la industria se han impuesto tanto a la concepción utilitarista como a la naturalista de los derechos de autor y de la propiedad intelectual. En la práctica, como veremos, la propiedad intelectual es utilizada para restringir la libre competencia y la innovación. En su estado actual, los derechos de propiedad intelectual son un estorbo para el desarrollo de la ciencia y de la cultura, para el acceso de los ciudadanos a la información y el conocimiento y se han convertido en una burla trágica de sus primitivos objetivos.

La propiedad intelectual, lejos de servir a sus objetivos iniciales, está siendo utilizada para facilitar el control de las grandes compañías sobre la cultura, el arte y la innovación tecnológica. De la misma forma, los derechos de autor, en lugar de incentivar a los verdaderos autores, que apenas se benefician de los mismos, están siendo utilizados para restringir las libertades de los individuos y permitir el mantenimiento del control de las grandes compañías sobre la creación y la distribución de libros, revistas, literatura, música, cine, software y demás” [Villate, 2001].

Por el contrario, esta preocupación de que los derechos de los usuarios no puede estar en desbalance con los derechos de los autores, ha sido recogida claramente por las bibliotecas y sus organizaciones. La Federación Internacional de Asociaciones e Instituciones Bibliotecarias (IFLA) establece claramente estas posturas en su documento *Position Paper on Copyright in the Electronic Environment* el cual fue sancionado por sus comités en 1996 en la 62ª Conferencia Mundial de esta organización. [International Federation of Library Associations and Institutions, 1996].

Como ha podido observarse, si bien algunas legislaciones han extendido las protecciones de sus leyes de derechos de autor mucho más allá de lo usual en materiales impresos, también puede observarse que existen ya algunos contrapesos que buscan emitirles a las bibliotecas continuar con su labor, conscientes que de otra forma corren el riesgo de ser asfixiadas por estas disposiciones. "...Por siglos hemos desarrollado y arraigado la idea de que si nos gusta algo que compramos para leer, podemos recomendarlo, prestarlo o regalarlo a un amigo o colega, y además si decidimos que ya no deseamos esa obra, podemos tirarla o venderla. Estos principios son tan importantes en nuestro quehacer intelectual que debemos tenerlos presentes siempre y encontrar la manera de trasladarlos al entorno de las publicaciones digitales" [Soules, A. 2002].

Vale la pena abundar un poco más en estas características *sociológicas* del libro. Conviene entonces remontarnos hasta los orígenes del mismo en su forma impresa. Antes de la imprenta de Gutenberg los libros sólo podían ser copiados a mano, uno a uno, y únicamente por las personas que pudiesen darse el caro lujo de tener el tiempo y la habilidad de copiar ese libro, o de tener los recursos para hacerlo copiar por otra persona. Estas personas no abundaban en esa época, y como era un proceso penoso, caro y artesanal, ningún derecho de autor o copia estaba involucrado. El copista podía tomar el todo o las partes: modificarlas, suprimirlas, comentarlas, iluminarlas, etcétera. En la práctica, cada ejemplar que se copiaba era una nueva edición. No era lógico prohibir o regular la copia de libros producidos exclusivamente en esa forma; nadie iba a lucrar con libros así producidos. Con el advenimiento del libro impreso por tipos, un cierto número de ejemplares eran idénticos dentro de esa edición. Si un artesano poseía una imprenta, las copias se podían mecanizar dentro del taller de ese impresor y era entonces fácil que esa persona copiara la obra de otro, con lo que obtendría múltiples copias y haría algún dinero con ello. En algún momento tuvo que legislarse a ese respecto. Aún así, pasaron nada menos que dos siglos y medio desde la aparición de la imprenta hasta que hizo su aparición la primera y tímida legislación autoral.

Nótese que las disposiciones que el *copyright* imponía eran limitaciones a acciones que ningún lector normal podía realizar por sí mismo,

por lo que de hecho no suponían ningún problema para el usuario. Los derechos de la propiedad intelectual intentaban regular productos tangibles, manufacturados. El fundamento estaba en el soporte físico. *El copyright era una regulación industrial*. En la medida en que se van conformando las legislaciones de derechos autorales a lo largo del siglo XIX es conveniente resaltar que el espíritu de la ley que animó la restricción para copiar en todas ellas nació como una restricción para otros impresores, no para el público. El derecho de copia, más allá de los derechos morales protege a un editor, a quien el autor le ha cedido los derechos patrimoniales de una edición, de otros editores, para que no le dupliquen su edición y le disminuyan sus posibles ganancias. En la reproducción vía imprenta el riesgo no provenía del público sino de otros impresores rivales; para el público no era práctico poder conseguir una imprenta y empezar a obtener copias de una obra. Por ello la venta de las copias ya impresas y pagadas por parte del público, no representaban ningún riesgo económico para el editor; ni siquiera si él decidía por algún medio manual obtener una copia para sí. No era necesario amenazar con cárcel a un lector que obtenía así una copia eventual.

Por lo mismo, el derecho de poseer físicamente un ejemplar de una obra por parte de un comprador fue consagrado en todas las legislaciones del mundo por casi dos siglos. Por ello podemos poseerlo indefinidamente, regalarlo, prestarlo, venderlo, subastarlo, anotarlo, transportarlo, heredarlo, leerlo en voz alta e inclusive destruirlo. Esto se ha vuelto un derecho intrínseco a la naturaleza del que posee un libro u obra semejante. Las bibliotecas y sus colecciones de libros o revistas encajaban perfectamente en este esquema —en realidad por él existen— y todo estaba en equilibrio.

Con el advenimiento de los mecanismos de copia masiva al alcance del público y de las bibliotecas este equilibrio empezó a romperse. La fotocopia en los impresos y la grabadora en los registros sonoros permitieron al público y a las bibliotecas obtener mucha mayor cantidad de páginas u obras. A partir de ese momento los editores arremetieron contra el derecho de copia, que había estado intrínsecamente ligado a un lector y que no había sido cuestionado en los siglos anteriores. Las bibliotecas pagaron caro el delito de *lesa copyright* por el

servicio que les daban a sus usuarios; los precios de las suscripciones y otros materiales se vieron incrementados para este tipo de institución, dado su nefando pecado de obtener copias para el público, y de algún modo el equilibrio volvió a ser reestablecido, a pesar de que, como ya hemos mencionado, las bibliotecas sólo aplican la esencia del derecho de copia consagrado en esa original premisa constitucional ratificada por la Suprema Corte de Justicia de los E.U.A. y contenida **en la mayoría** de las legislaciones autorales del mundo: “el propósito del ‘copyright’ es promover la ciencia y las artes útiles a través de la difusión del conocimiento” y que “el autor se beneficie de la distribución ya que ello disemina su obra, creando además de nuevo conocimiento, nuevas ventas”.

Con el advenimiento de las publicaciones electrónicas y la posibilidad de las bibliotecas y del público de efectuar copias muy fidedignas y masivas de las obras, es necesario poner en contexto nuevamente las legislaciones mundiales para que el equilibrio sea restablecido. Es cierto que el público se vuelve ahora un riesgo potencial para el editor en la medida que puede copiar y distribuir masivamente una obra y por ende debe ser controlado, pero también es verdad que el público no puede ser tratado y restringido partiendo del supuesto de que **siempre** se convertirá en otro editor comercial. El público ha creado derechos y costumbres de cómo usar una obra y a obtener sus “copias incidentales temporales”, —como ha sido ratificado nuevamente en la DMCA— a lo largo de siglos, y es imposible ahora simplemente pretender que cuando se paga por una obra sólo se adquiere el derecho de ver esa obra por un periodo finito de tiempo, o en un sólo lugar, y que es exclusivamente para sus ojos y no puede ser compartida en modo alguno. Eso va en contra de la naturaleza hombre-libro. Ningún modelo comercial, tecnológico o legal de alta restricción entre los establecidos a la fecha parece prometer algo real para el futuro. El equilibrio entre el derecho de comercializar de unos y el derecho de copiar de otros debe ser reestablecido en la era digital; la fórmula debe ser de ganancia para ambos; ninguna otra fórmula funcionará a la larga. Como afirma la sabiduría popular: *Todos los extremos son malos*.

Citando nuevamente a Villate:

“Hay quienes dicen que el ciberespacio no solo ha modificado la tecnología de copiado, sino también el poder de la ley para luchar contra la copia ilegal. Internet y el entorno digital no sólo permiten hacer copias idénticas a los originales y distribuirlas masivamente con costes prácticamente nulos, sino que, además, amenazan con imponer una tarea casi imposible a los encargados de velar por el cumplimiento de la ley: localizar y castigar a quienes cometen violaciones sobre los derechos de autor. La amenaza de la tecnología contra los derechos de autor sería máxima, mientras que la protección que la ley puede ofrecer a esos mismos derechos sería mínima.

Ésta es la visión de las cosas que ha conseguido imponer la industria de los contenidos y sus grupos de presión. ‘Internet es el reino de la anarquía y de la piratería; y si no se pone orden, editores, discográficas y estudios cinematográficos no tendrán incentivo alguno para derramar sus contenidos en el ciberespacio. Y si esto no sucede, Internet no podrá ser ese fondo inagotable de información y cultura prometido’.

Este discurso ha encontrado oídos receptivos. Las grandes compañías discográficas, cinematográficas y editoriales son grupos muy poderosos en todos los países del mundo. Los informes sobre “piratería” y sobre las consiguientes pérdidas multimillonarias de la industria llegan con extrema facilidad a los medios de comunicación y a los despachos de políticos y legisladores. El temor y la ignorancia hacen el resto.

En estos momentos, tanto a nivel internacional, como regional y estatal se han impuesto legislaciones draconianas que respaldan buena parte de las desmedidas pretensiones de la industria. Así ha sucedido en los nuevos tratados de la Organización Mundial de la Propiedad Intelectual (OMPI), y en la nefasta ley DMCA americana”. [Villate, 2001].

Los editores impulsaron de origen el *copyright* como una medida para protegerse de sus rivales, los otros editores; ha sido y es un grave error ver al público como un rival o un enemigo a vencer. La “piratería”, es decir, la aparición de copias ilegales de obras es causada, entre otras cosas, por los editores y aprovechada por vividores cuando los editores tratan ellos mismos con actitudes de “pirata” a su público usuario: extensiones una y otra vez de derechos de explotación de obras más que explotadas y de las cuales ya se obtuvieron pingües ganancias;

diferencias cada vez mayores e injustas entre el costo de producir, que tiende a reducirse, y el precio final al usuario, que tiende a incrementarse; cadenas de distribución enormes e innecesarias en donde demasiados participantes obtienen ganancias a costa del usuario; modelos de uso cada vez más restrictivos de las libertades del usuario final, obsolescencia forzada de tecnología con modelos de recompra obligada del mismo material por parte del usuario, etcétera.

A la fecha las soluciones planteadas al problema generado por la facilidad de copia y distribución de la información digital han atacado más a los síntomas que a las causas verdaderas del problema. El simple endurecimiento de las leyes y las medidas coercitivas o las trabas tecnológicas no ha impedido que las copias ilegales proliferen en todo el mundo explosivamente, y en algunos casos en efecto dañan ya las legítimas ganancias de los editores ¿no es tiempo ya de ir buscando otro modelo que restablezca *efectivamente* el equilibrio entre el editor y el usuario dentro del contexto digital? El desprecio de los editores al principio esencial histórico del derecho de copia es sin duda una de las principales razones de la falta de acuerdos en este contexto actual y lo que ha producido en mayor escala es el desprecio del público hacia sus derechos de comercialización, factor que aprovechan otros editores ilegales sin escrúpulos para crear la industria de la “piratería” de obras.

En resumen, pienso que si los editores rescatan los principios y el espíritu básicos que han regido a los derechos de copia y los modelos de negocio de algunas industrias editoriales, no tan sólo la del libro, y tales derechos son replanteados con honestidad podrían encontrarse mejores soluciones para el mundo digital; algunos de ellos ya se han mencionado y vienen de origen desde la primera legislación autoral: los autores y editores deben protegerse de otros editores sin escrúpulos que lesionan sus intereses; ellos son los verdaderos antagonistas y tienen todo el derecho de luchar contra ellos. Sólo cuando un miembro del público decide convertirse en “editor” ilegal se vuelve un antagonista: considerar al público en general siempre como otro editor potencial restringiéndole absolutamente sus prerrogativas de copia “legal” o ética va contra la esencia misma que creó el libro y contra la generación y la divulgación del conocimiento. Ningún modelo que

vaya en contra del restablecimiento del balance entre el legítimo derecho de copiar y poseer de los usuarios y el legítimo derecho de comercializar y obtener ganancias de los autores y editoriales funcionará. Cualquier cosa orientada hacia uno de los extremos estará destinada al fracaso escindirán cada vez más el espacio entre editores y usuarios, con el beneplácito de quienes detentan esas obras para obtener un lucro indebido. Usuarios y editores deben plantearse mutuamente el paradigma: *si tú me ves como tu socio, seré tu socio; si tú me ves como tu enemigo, seré tu enemigo*.

Como corolario a esta problemática, presento una pequeña analogía de un libro de papel con un libro electrónico o un DVD, con el fin de ilustrar los absurdos a los que estamos llegando en la búsqueda precipitada de soluciones al problema sin atender a sus verdaderas causas y efectos. La idea para construir esta analogía fue tomada a partir del comentario que hizo un usuario de una lista de discusión electrónica sobre un sitio llamado *Slashdot* y cuya única referencia es la clave *DarkbladePDX*.

“Imagine que usted realiza una estancia en España. Visitando una librería encuentra un libro cuyo título promete que la obra es interesante y decide comprarlo; lo lleva a casa, rompe la envoltura y lo abre. En la parte interior de la cubierta delantera del libro se observa un sello que indica en qué parte del mundo compró usted el libro: España. Examinada la cubierta llega a la portada y después de ella, usted encuentra una página con el número 11. Después, al comenzar a leer el libro, en vez de castellano claro, descubre un trabalenguas: *Rzxs Dzse Bfs Wlv jAoxdo*; todo el resto del libro se observa igual. Después de examinarlo un poco más a fondo, y como usted tiene cierta intuición para la criptología, ¡ahá!, descubre que el libro sí está impreso en español, pero con un encriptado simple aplicando algo tan sencillo como mover cada letra hacia adelante ‘n’ posiciones en el alfabeto, donde ‘n’ es el número que aparece en la página posterior a la portada: once. No obstante lo simple del encriptado, debido a éste, sin descifrarlo usted no puede leerlo.

Algo sorprendido por el hecho, usted lleva el libro de vuelta a la librería. Ellos le dicen que sobre la caja había un letrero que decía: *salida*

la mercancía no se admite reclamación, y que ellos sólo pueden cambiar un título por otro título idéntico, y que *de todos modos Juan te llamas* ya que todas las copias del libro están codificadas. No obstante, y con un genuino y desinteresado afán de servirle pueden alquilarle a usted temporalmente a un empleado del editor apodado ‘el decodificador’ para llevarlo a vivir a su casa —la de usted, no la de él— y que él se encargue de leerle los libros encriptados, ya que este decodificador está perfectamente entrenado para leer ese código. Como usted quiere leer el libro, no tiene más remedio que aceptar, pagar y llevarlo a casa. Ya ahí, usted le da indicaciones de que lea el libro; él busca el número clave en el punto particular de la página de números y hace al instante el corrimiento de caracteres del alfabeto para el cual él está entrenado perfectamente, sea cual sea el número clave y lee así el libro mientras usted lo desee. Parece que sí hubo una solución. Usted sigue comprando libros bajo ese esquema y el decodificador se los lee. Sin embargo, hay un *pero* molesto: el decodificador insiste en que por políticas de su editorial es obligatorio que él lea siempre la introducción enteramente cada vez que se le pida que comience a leer cualquier parte del libro, la cual consiste mayormente en publicidad para otros libros del editor. ¡qué se le va a hacer! Adelante.

Usted regresa de su viaje a su casa en México y observa que acumuló una buena cantidad de libros. Por supuesto, usted no se trajo al decodificador de España pues sabe que puede contratar uno en México. Va a la librería y de ahí trae a su decodificador así contratado para poder seguir leyendo sus libros. Una vez en casa le pide a éste que se los lea; pero entonces descubre que el decodificador tiene instrucciones precisas de su empresa y debe rehusarse a descifrar otros libros que no provengan de su parte natal del mundo, aún siendo de la misma editorial. Como el decodificador actual es mexicano, sólo puede leer libros cuyo sello de compra sea “México y Sudamérica”, y por lo tanto no descifrára los libros comprados en España, y usted tendrá que ir a la librería a alquilar otro decodificador, nativo de ese país. Esto puede volverse muy caro y engorroso si usted tiene la mala costumbre de comprar libros de todo el mundo. Además, el editor se reserva el derecho de despedir al decodificador en cualquier momento sin reembolsarle

a usted su dinero, y el decodificador no será capaz de descifrar nuevos libros porque el editor no lo autorizará a hacerlo para nadie más que para él.

Como consecuencia, usted posee un libro que ha sido hecho de tal forma que usted mismo no pueda leerlo, y la única persona a la que le permiten que lea para usted no puede ser contratada permanentemente por usted bajo sus condiciones; sólo se la alquila temporalmente, y el pago de sus servicios no será en modo alguno reembolsable. Como usted es un lector con iniciativa, y no piensa renunciar a sus libros ***por los cuales ya pagó***, se plantea llevar a cabo alguna de las siguientes soluciones:

- Aprender usted mismo a descifrar y leer libros codificados.
- Enseñarle a alguien cómo descifrar los libros para que se los lea.
- Grabar o hacer una transcripción mientras su decodificador lee el libro.
- Alquilar otro decodificador que no trabaje para ese editor.
- Convencer al decodificador a que descifre libros externos a la región natal de éste.
- Hacer que el decodificador se salte la introducción.

Sólo que, de acuerdo con el Título 17, Capítulo 12, Sección 1201 del “Acta de Copyright del Milenio” —DMCA— de los E.U.A., es absolutamente ilegal hacer cualquiera de esas cosas, bajo severas penas de multas y cárcel.

Sumamente contrariado, usted decide llamar a la editorial quejándose de esta molesta situación. En la editorial le contestan muy amablemente: que sí, que están conscientes, pero que son las nuevas políticas de la empresa para proteger sus intereses contra la piratería. Empero, en un legítimo y sincero afán de servir a sus lectores ya están poniendo a la venta nuevas modalidades de libros como contraparte a la modalidad del decodificador, la cual, recién han descubierto, no es bien aceptada por el público. Ahora, y para su comodidad y beneplácito, tienen libros con otras características muy *modernas* y que no tienen encriptado: hay unos escritos en español normal, con la única

limitante de que sus páginas se oscurecen después de diez días de adquirido y resulta imposible leerlo más. Si no quiere el límite de días hay otra versión del libro cuyas páginas se oscurecen un poco cada vez que se leen, hasta que después de un cierto número de lecturas, queda ilegible. También le ofrecen otros libros que no tienen esas limitantes, siempre y cuando se acepte y se jure solemnemente que el libro será leído siempre en la habitación en que se abra por primera vez, pues no es posible leerlo en ninguna otra. ¡Ah!, y no está permitido prestárselo o regalárselo a un amigo. Lamentablemente —para usted, no para ellos—, si usted desea comprar estas nuevas versiones de los libros, deberá pagar nuevamente por esos ejemplares, ya que *por desgracia* no se contempló ningún mecanismo de canje para los que ya tenía”.

Por absurda que esta analogía con un libro impreso pueda sonar, todas son prácticas comerciales ya establecidas por empresas editoriales de DVD con películas, libros electrónicos, revistas, etcétera. Y las restricciones de la ley DMCA son absoluta y totalmente reales, en lo relativo a tecnologías anti-copiado que se encuentren embebidas en los materiales electrónicos. El “decodificador” descrito aquí sí existe: es en la vida real un “*software*” embebido en películas dentro de DVD que están a la venta en la mayor parte del mundo. En los Estados Unidos y en otros países del orbe, es ilegal modificar una tornamesa DVD para que reproduzca discos de otras regiones del planeta, o para que se brinque la introducción, o para que ignore las técnicas para evitar la copia a otro dispositivo. Es incluso ilegal divulgar cómo funciona el *software* o cómo ignorarlo. Tal vez se piense que esto sólo es un problema de los DVD, pero lo peor es que estos usos, costumbres y leyes ya empiezan a ser observados con alarmante y molesta frecuencia en dispositivos reproductores y productos de libros electrónicos. Las variantes de libros o artículos que se “desvanecen” después de cierto tiempo *ya* existen y se ofrecen a la venta en libros electrónicos. Muchas de las versiones “descargables” de textos electrónicos sólo pueden ser instaladas en una sola computadora por una sola vez. No es posible trasladar la obra de la casa a la oficina, o prestársela o regalársela a otra persona, y si la computadora sufre un daño es prácticamente imposible recuperar la instalación por parte de la editorial. Puede ser que estas “soluciones” resuelvan el problema del editor, pero menosprecian y agreden de la

forma más absurda siglos de una relación establecida hombre-libro por parte del usuario. ¿Es éste el prometedor futuro de la publicación electrónica? Nótese que no se trata de un problema de falta de capacidad tecnológica, sino de un problema de las restricciones tecnológicas que hemos agregado a falta de *soluciones sociales de fondo*. Sin duda somos capaces de idear algo mejor.

Durante el Claustro Pleno del 11 de marzo de 1776 de la Universidad de Salamanca,⁷³ el primer reglamento de bibliotecas conocido en España, estableció en su capítulo sexto: “De la custodia de los libros y su reponsabilidad. La experienzia ha demostrado que no están de más, y que antes bien son utilíssimas todas las precauciones que se toman para la más segura custodia de los libros, y siendo una dellas la de fijar zensura en las puertas principales de la Bibliotheca. Se tomará esta probidenza para que ninguna persona de qualesquiera condizi3n o calidad pueda, sin espreso mandato de la Vniversidad, o de alguna de sus junctas de facultades, o librería, extraher libro alguno de la Bibliotheca, papel ni alajas della, sin incurrir en excomuni3n maior...”(sic). Sería bueno recordar que ni libros encadenados ni excomuniones en la puerta de las bibliotecas ni anatemas perduraron; no iban al fondo y no resolvieron el problema.

Mientras que las maldiciones en los libros eran poco efectivas contra los bibliómanos, se ha establecido que tuvieron además un efecto negativo en la forma de compartir información durante la sociedad medieval. Bajo el pretexto de la excomuni3n o el anatema, muchos monasterios rehusaban prestar sus libros y este hábito se extendió a la sociedad seglar. Dada la frustraci3n debida a la falta de materiales para los estudiantes de la Universidad de la Sorbona, el Concilio de París de 1212 declaró: “Prohibimos a aquellos pertenecientes a una orden religiosa formular cualquier voto contra el préstamo de sus libros a aquellos que los necesitan [...] desde esta fecha, ningún libro podrá ser retenido so pretexto de excomuni3n u otra pena [...] Después de las debidas consideraciones, dejen que algunos libros permanezcan en la casa para el uso de los hermanos; pero permitan que otros, de

73 Archivo de la Universidad de Salamanca (AUS). Libros de Claustros, Libro 239, marzo 11, 1776.

acuerdo a las decisiones del abad, sean prestados a quienes están en necesidad de ellos, salvaguardando los derechos de la casa [...] y anulamos y otorgamos absolución para todos los anatemas de este tipo y esas penas quedan sin efecto”.⁷⁴ Esta declaración aparenta ser una declaración fuerte contra la posición típica de la iglesia acerca del préstamo de libros e ilustra el grado de consecuencias negativas que ciertas medidas aplicadas con exceso pueden causar en la forma de compartir información. Hubo desde entonces quien estuvo consciente de que la universidad no podría producir doctores, abogados, etcétera, necesarios para el desarrollo adecuado de la sociedad si la transferencia de información era asfixiada bajo cualquier pretexto.



“Palacio de Asurbanípal, rey del mundo, rey de Asiria, que confía en Asur y en Ninlil, y a quien Nabu y Tashmetu proporcionaron finos oídos y perspicacia: Lo mejor del arte del escriba, que ninguno de mis antecesores consignó, la sabiduría de Nabu. Los signos de la escritura; todos los que han sido inventados los he escrito en tabletas, las he ordenado en series, las he colacionado y las he colocado en mi palacio para mi real contemplación y lectura. Tu señorío no tiene igual, Asur, rey de los dioses. Cualquiera que remueva la tableta y ponga su nombre junto al mío, derribenlo, contrariados y enfadados Asur y Ninlil, y hagan desaparecer su nombre y su especie de la tierra”. (circa 650 a.C.)

⁷⁴ Delisle, Léopold . 1848. en “Bibliothèque, de l'Ecole des Chartes”, ser. 3, I, p. 225. Citado por Thompson, James W. 1967. *The Medieval Library*. New York: Hafner Publishing.

5.2. INICIATIVAS INTERNACIONALES SOBRE LOS DERECHOS

Considero relevante destacar aquí dos documentos importantes acerca del tema de los derechos de propiedad. Por un lado se encuentra el documento de la Unión Europea, en donde puede observarse una visión de nación y de región muy adecuada para la planeación de decisiones estratégicas y de largo alcance. Por el otro está el documento, cada vez más difundido, acerca del *fair use*, uso ético de la información, el cual le permite al administrador de una biblioteca o conjunto de ellas tener una directriz al interior de su institución mientras se legisla con más precisión a propósito de la información digital.

5.2.1 La iniciativa europea

La versión más avanzada de la Unión Europea con respecto a cómo se observa el futuro cercano de los derechos de propiedad se encuentra en el documento *Propuesta para los Derechos Afines en la Sociedad de la Información COM (97) 628] final COD97/0359* [European Commission, 1997. Com. 628]. La Comisión, basándose en las contribuciones de los sectores interesados, ha determinado nueve temas en su opinión prioritarios para establecer los regímenes de protección de los derechos de autor y derechos afines de cara al funcionamiento de la Sociedad de la Información. La Comisión recabó durante tres años la opinión de los sectores interesados sobre los aspectos tanto técnicos como normativos. Resumiendo, las ideas más relevantes de esa propuesta son:

- El éxito del proceso de desarrollo de la Sociedad de la Información depende de que la mayoría de los servicios y productos nuevos que se creen puedan beneficiarse plenamente de las redes de información. Su pleno desarrollo ha de basarse en un marco normativo coherente en el ámbito nacional, comunitario e internacional, simultáneamente. Será sin duda necesaria la adaptación de ciertas normativas para responder a las nuevas y diferentes exigencias que puedan surgir y plantearnos nuevos desafíos. La adaptación de los regímenes jurídicos relativos a la propiedad

intelectual constituye uno de ellos. El enfoque que ofrece la legislación relativa al mercado interior esboza las líneas maestras para elaborar las políticas sobre la sociedad de la información.

- Estos servicios y productos nuevos, que se difundirán a través de las redes de información, o bien recurrirán a obras preexistentes, o bien conducirán a la creación de obras nuevas. Las obras existentes deberán adaptarse frecuentemente antes de su transmisión a través de un entorno digital, y resulta obvio que la creación de obras y servicios nuevos implicará cuantiosas inversiones, sin las cuales el contenido de los nuevos servicios que circulen en la Sociedad de la Información sería muy limitado.
- Los esfuerzos creativos para promover las inversiones en los nuevos servicios sólo podrán realizarse y justificarse si en este entorno digital la protección conferida por los derechos de autor y los derechos afines a las obras y prestaciones resulta suficiente. Una vez puesta en servicio en la red es muy difícil, sin la adecuada protección, garantizar que la obra o la prestación no sea copiada, transformada o explotada sin el conocimiento de los derechohabientes y en su perjuicio.
- Por las propias características de las redes que operan en la sociedad de la información, la existencia de grandes diferencias en los grados de protección de las obras y las prestaciones acarrearía obstáculos a la creación de la sociedad de la información. En efecto, la dificultad de controlar la utilización de la obra y las posibilidades de reducir lo local de los agentes económicos que ello conlleva hacen necesaria, en algunos ámbitos, una mayor armonización de la protección conferida por los derechos de autor y los derechos afines.
- A escala comunitaria, ya se ha alcanzado cierto grado de armonización, gracias a la adopción de cuatro directivas relativas a los derechos de autor y derechos afines. Este marco jurídico se completará probablemente en breve con la directiva relativa a la protección jurídica de las bases de datos. Esta medida sitúa a la comunidad muy por delante de sus socios comerciales, en la medida en que le permite dotarse de un marco jurídico que sirva de apoyo al desarrollo de los servicios de la Sociedad de la información.

- De no llevarse a cabo la armonización de las normas en materia de derechos exclusivos de autorización o de prohibición de la utilización, la reproducción, etcétera, de las obras y prestaciones protegidas, la libre circulación de los productos y servicios concernidos se vería obstaculizada. Ésta es la razón por la cual sólo mediante la armonización de las legislaciones de los estados miembros podrá restringirse la delimitación territorial de los derechos concedidos en virtud del derecho nacional. Por otra parte, sin el suficiente grado de armonización a escala comunitaria, los mercados que se abran a los nuevos servicios podrían permanecer fraccionados, de impedir así el desarrollo de servicios que requieren una explotación en un mercado más amplio que el nacional para ser rentables.
- Por consiguiente, deben abordarse a la brevedad varios asuntos verticales, los problemas relativos a ciertos derechos específicos y otros relacionados con la explotación de los derechos. En cuanto a los aspectos horizontales, se trata del derecho aplicable y del agotamiento de ciertos derechos. Las cuestiones relativas a los derechos específicos se refieren al derecho de reproducción, al concepto de público en el ámbito del derecho de comunicación al público, y al análisis de determinados derechos específicos que podrían otorgársele a los diferentes tipos de transmisiones digitales. El problema del derecho moral será también objeto de un examen pormenorizado. Por último, en la explotación de los derechos deben analizarse los problemas relativos a la gestión de los derechos y los derivados de los sistemas de identificación y de protección desde un punto de vista técnico.

De todos los documentos que he revisado, éste de la Unión Europea me parece, si no el mejor, *el menos malo*, al ser el más avanzado y balanceado en los aspectos que hay que considerar desde un punto de vista estratégico en relación con los derechos de propiedad y con los de la Sociedad de la Información y tiene un enfoque de largo plazo. Si bien para mi gusto todavía sigue cargado hacia el lado de privilegiar a los editores, habrá que estar muy atento a sus avances y desarrollo.

5.2.2 El *fair use* o uso ético de la información.

No es necesario abundar acerca de la velocidad y variedad del desarrollo que ha alcanzado la información electrónica en la última década. Debido a esta explosión, mucha de la legislación establecida a nivel mundial no contempla casos específicos para la producción editorial electrónica. Ante la falta de legislaciones o la existencia de lagunas a propósito de la información digital que espera las consiguientes modificaciones a las legislaciones existentes, y mientras esto va sucediendo, poco a poco en diversos países, se ha difundido en años recientes un código de ética que ha servido cada vez más como guía para muchos bibliotecarios del mundo que tienen que distribuir información digital. Como ya hemos mencionado, una de las legislaciones más restrictivas en los aspectos de copia es la legislación estadounidense, la que en su afán de proteger a los creadores de obras terminó sobreprotegiéndolos a costa de los usuarios y metiéndose en aprietos ante el advenimiento de la información electrónica con fines académicos. Para fines prácticos, todo es ilegal bajo la legislación norteamericana; el problema se agudiza por el hecho de ser ése un país donde las demandas legales por cualquier minucia son el pan de todos los días y han tenido que encontrarle *interpretaciones* a esa normatividad. De ello observamos como producto las recomendaciones del “uso ético”, las que como su nombre indica, son guías sin carácter obligatorio, pero consensadas y aceptadas tácitamente por la generalidad del medio estadounidense, y con el tiempo se han convertido en una guía muy útil en otros países en donde la falta de normatividad crea dudas entre los bibliotecarios y los académicos. Para el individuo administrador de bibliotecas éste es un código muy útil.

Una introducción breve pero concisa puede encontrarse en un memorándum que les envió en 1998 la célebre Condoleezza Rice,⁷⁵ en su calidad de Preboste de la Universidad de Stanford, a los miembros de la facultad; en él, establece algunos de los antecedentes y las características

⁷⁵ Rice, Condoleezza. 1998. *Copyright & Fair Use. Memorandum to Faculty Members*. Stanford University Libraries. Disponible noviembre 3, 2005 en: http://fairuse.stanford.edu/stanford_notices/rice.html

aceptadas para el *uso ético* de la información pensando en las aplicaciones de docencia e investigación de esa universidad. Entre ellos se destaca el hecho de que con este enfoque se permite la reproducción limitada de obras protegidas con fines de investigación, crítica, reportaje, parodia y docencia —incluyendo varias copias para uso en el aula—. Los cuatro principales parámetros para determinar si un documento dado se copia bajo el principio del *uso ético* sin infringir la ley son:

- *El propósito y carácter del uso, considerando si éstos son de naturaleza comercial o tiene fines educativos no lucrativos.*

¿Es ésta una copia igual al original? si lo es, no parecería estar cumpliendo este criterio. ¿la nueva obra ofrece algo más allá del original o lo transforma sustancialmente? este caso parece estar más acorde con el uso ético. ¿El uso pretendido es sin fines de lucro? ¿o tiene fines académicos? esto parecería estar acorde con el ‘uso ético’.

- *La naturaleza del trabajo protegido.*

¿La obra ha sido publicada o no? Las obras no publicadas generalmente no entran en el uso ético, y no por estar sin publicar carecen de derechos de autor. Las obras descatalogadas normalmente sí entran aquí.

¿La obra es factual o artística? Entre más tiende una obra a ser una expresión artística, menos parece caer su copia dentro del ámbito del uso ético.

- *La cantidad y relevancia de la porción copiada, en relación con la obra en su conjunto.*

Entre menos parte de la obra se tome, existe menos riesgo de que salga del ‘uso ético’. Debe tenerse particular cuidado cuando la parte que se toma es la ‘esencia’ de la obra.

- *El efecto de ese uso sobre el mercado o valor potencial de la obra copiada.*

¿Qué tanto esta copia produce un efecto negativo en las utilidades del titular de los derechos?

Aunque todos estos factores se considerarán, el último factor será el más importante para determinar si cierto uso es *ético*. Ahí donde una obra se encuentre a la mano por medio de compra o licenciamiento por parte del dueño del derecho en el medio o formato deseado, copiar el todo o una porción significativa del trabajo en vez de comprar o licenciar un número suficiente de copias “autorizadas” habría presuntamente un comportamiento *antiético*. Donde sólo una porción pequeña de la obra fuera copiada y ésta no se utilizara mayormente y se hubiera comprado o licenciado un número suficiente de copias autorizadas, el uso de esa a copia sería con gran certeza *ético*.

Si se observa con cuidado, puede notarse que los tres primeros de esos cuatro aspectos están perfectamente cubiertos por el artículo 148 de la ley mexicana al efecto, y con carácter de ley, no de recomendación. Por ello, agregando el cuarto criterio a los siete apartados de nuestra ley, podemos tener una guía muy consistente acerca de lo que puede ser copiado y lo que no. Nótese que se afirma “muy consistente”. No hay absolutamente ninguna receta que permita realizar una copia cien por ciento exenta del riesgo de una reclamación, sobre todo en el ámbito internacional, pero los criterios del *uso ético* tienden a ser más y más utilizados como un estándar *de facto* y su ventaja enorme es su aceptación internacional. Por ello, el número de bibliotecarios que se rigen por él es cada vez mayor en el medio de la información digital.

Complementando estas reglas se encuentran los llamados *criterios para maestros e instructores*, los *criterios para alumnos* y los *criterios para las escuelas*, los cuales en esencia se rigen por las siguientes recomendaciones [University of Maryland, 2000]:

Para los maestros o instructores:

- *Criterio de brevedad*; es decir, que sólo se copian partes mínimas del total de una obra: un capítulo de un libro, y no todo el libro; un artículo de una revista o periódico; una historia corta,

ensayo, cuento o poema dentro de una antología. Algunos diagramas, tablas, dibujos, etcétera, pero no la cantidad total de los que aparecen en la obra, etcétera.

- *Criterio de espontaneidad*, se relaciona con la frecuencia con la que se copia una obra. Si se copia para un curso en una escuela dada, puede considerarse uso ético. Si se copia y se usa a lo largo de semestres y semestres o para un número considerable de cursos o escuelas esto se aleja del criterio mencionado.

Para los alumnos:

- *Criterio de igualdad*; establecido, que deben respetarse los derechos de autor independientemente de que el material se encuentre en un medio tangible tradicional —papel, disco, video, etcétera—, como en un medio electrónico— base de datos, cd-rom, boletín electrónico, o Internet—.
- *Criterio de uso personal*. Una copia de un documento para el estudio propio conlleva un uso ético. Reproducirlo o redistribuirlo en un sitio *web* del alumno probablemente no lo sea. El alumno puede incorporar porciones de material tomado de otras obras para proyectos o investigaciones de sus cursos o graduación, para portafolios personales, para solicitudes a otras escuelas y/o grados o a un empleo. Pero debe descartarse cualquier aplicación comercial.

Para las escuelas:

- *Criterio de aplicación curricular*. La escuela puede tomar porciones de material para elaborar materiales de apoyo en proyectos multimediales para sus programas de enseñanza curricular. Este material puede ser usado para el estudio de los alumnos o para instrucción remota siempre y cuando el acceso sea restringido y seguro.
- También puede ser utilizado en conferencias, presentaciones, seminarios o talleres. La temporalidad de estos materiales no debe exceder los dos años sin obtener los permisos correspondientes.

En todos los casos siempre debe darse en las copias claramente el crédito correspondiente de la autoría de la obra, y seguir las instrucciones de los autores cuando las haya.

La versión más útil de estas guías es la más actualizada, completa y, de algún modo, “oficial”, en el sentido de que está consensada por un importante número de instancias, y puede encontrarse en las *Guías para el uso ético de multimedios educacionales establecidas por el Congreso de los EUA*, la misma instancia que aprobó la malhadada DMCA. (*Fair Use Guidelines for Educational Multimedia established by Congress of the United States: Nonlegislative Report of the Subcommittee on Courts and Intellectual Property Committee on the Judiciary US House of Representatives*). [United States of America. House of Congress. 1996]. En estas guías se detallan con más precisión el significado de algunos de los elementos como “cantidad y relevancia de la porción copiada”, etcétera. Como está diseñada para propósitos educativos *y está pensada y consensada ya* para materiales digitales representa una gran ayuda para el bibliotecario en estos tiempos de indefiniciones, a pesar de que, como ya se ha mencionado, no tienen el carácter de ley; recordemos que hablamos de “uso ético” el cual no es sinónimo de “uso legal”. Afortunadamente en algunos medios los guías empiezan a ser consideradas un estándar *de facto*. Para dar una idea de la relevancia que empiezan a tener, es conveniente mencionar la importancia de algunas de las instancias que sancionaron ese documento; entre las principales destacan:

Association of American Publishers (AAP) que incluye a más de 200 editores

Association of College & Research Libraries (ACRL)

Association for Educational Communications & Technology

Association for Information Media and Equipment (AIME)

Association of Research Libraries (ARL)

American Library Association (ALA)

American Society of Journalists and Authors (ASJA)

American Society of Media Photographers, Inc. (ASMP)

American Society of Composers, Authors and Publishers (ASCAP)

Association of American Colleges and Universities (AAC&U)

Business Software Alliance

John Wiley & Sons, Inc.

Information Industry Association (IIA); incluye 550 empresas.

McGraw-Hill

Microsoft Corporation

Motion Picture Association of America (MPAA)

Software Publishers Association (SPA); incluye 1200 editores de *software*.

Special Libraries Association (SLA)

Time Warner, Inc

U.S. National Endowment for the Arts (NEA)

U.S. Copyright Office

U.S. Patent and Trademark Office.

5.2.3 El *Copyleft* y *Creative Commons*.

El concepto del *Copyleft* tiene su origen en el sistema de comercialización o uso de los programas de computadora —*software*—. Bajo ese esquema típico, el usuario no adquiere la propiedad del programa, sino sólo el derecho a usarlo; por lo mismo, no puede mejorarlo, modificarlo o adaptarlo en forma alguna; es más, de hecho en la mayoría de los casos se prohíbe ver hasta el código fuente que lo origina. A este esquema se le denomina de *software propietario*.

Como contraposición a él, desde hace varios años se creó un esquema de *software libre* bajo el cual numerosos autores de piezas de programación ponían a la vista y bajo el dominio público sus programas permitían no sólo su uso, sino su distribución y modificación. Pero, ¿cómo evitar que alguna persona o institución, al hacer algunos cambios no decidiese registrar toda la obra a su nombre y ponerla bajo el regimen de protección del *copyright*, con lo cual evitaría que esta obra siguiese su camino en el dominio público? La respuesta fue el *copyleft*, concepto popularizado por Richard Stallman, presidente de la *Freeware Foundation* y, principal promotor mundial del concepto de *software libre*. El propósito del concepto del *copyright* es

el de mantener privado y bajo un propietario el uso de una obra, en este caso del *software*.

Como contrapeso, y de ahí su nombre, el propósito del *copyleft* es el de mantener libre a una pieza de *software* dada, y evitar así su privatización. Bajo este principio la declaración del *copyleft* en una obra es, en apariencia, la misma que la del *copyright*, e informa sobre el nombre del autor y sobre su calidad como creador o sobre propiedad intelectual de la obra, sólo que en el *copyleft* el autor autoriza explícitamente a los usuarios a agregar, modificar, mejorar, adaptar, etcétera, los trozos de código que considere pertinentes. Permite además obtener copias de la obra y distribuirlas libremente a discreción, pero lo más importante de todo, es que agrega una restricción, ***la de agregarle restricciones propias*** a los subsecuentes usuarios o autores. Esto es de singular importancia, ya que esta restricción de restringir, paradójicamente, se vuelve la clave de una distribución libre. De esta forma, cada usuario es libre de distribuir la obra, pero no puede imponer restricciones adicionales a su distribución. Ese nuevo usuario puede modificarla, señalando incluso dentro de la obra la parte que es de su autoría, pero no puede restringir la capacidad de otros para a su vez modificarla o distribuirla. De esta manera se garantiza que la obra no podrá ser llevada al régimen propietario más adelante por alguna persona o institución malintencionada.

Si bien el concepto del *copyleft* ha sido utilizado como un mecanismo de distribución de *software*, ha probado ya sus bondades y ha ganado muchos adeptos; por ello ha empezado a extenderse a otro tipo de obras, como las musicales o a los textos, sobre todo entre aquellos que viajan a través de la red mundial. Cada vez es más frecuente encontrarse con algún tipo de documento u obra de texto o artística que establece su esencia de *copyleft*, con lo cual mantiene su capacidad para circular ampliamente y sin restricciones por la red.

Una última aclaración: el concepto de *software libre* no atenta por sí mismo, como muchos predicán, contra las intenciones de obtener ganancias monetarias de una obra. No es un concepto de creación con fines sólo altruistas. *Software libre no implica tácitamente software gratuito*; el autor puede comercializarlo y pedir y obtener ganancias sobre su creación. En español es muy clara la diferencia entre *libre* y

gratuito. En inglés el término ha creado muchas confusiones y rechazos de origen ya que *free software* implica indistintamente *software libre* y *software gratuito*, haciendo que muchos posibles creadores lo consideren un concepto que sólo es para ser usado con intenciones altruistas, pero que no puede ser considerado cuando el propósito es el de obtener ganancias. Nosotros, gracias a nuestra herencia cervantina, podemos captar más fácilmente la diferencia, que sí existe, y evitar la confusión y el rechazo *a priori*. Por supuesto la misma conceptualización puede extenderse a la creación de obras de texto, artísticas, etcétera, y bajo este concepto podemos decidir si queremos ceder nuestra obra libremente, pero mediante un pago, o libremente y además gratuitamente.

Finalmente, como idea derivada de estos conceptos del *uso ético* y del *copyleft* encontramos otra instancia interesante denominada *Creative Commons*. En esencia, se trata de una metodología que permite al público en general que desea publicar una obra en *Internet* —texto, música, imagen, video, etcétera— seleccionar al instante en su sitio *Web* y sin ayuda de asesoría legal, una serie de diversas *licencias tipo* que ya han sido redactadas por expertos en donde el autor selecciona ciertos derechos que él quiere otorgar a quienes lo consulten, y ciertos derechos que él se reserva. Las combinaciones de derechos son muy numerosas y posibilitan que el autor pueda hacer una cesión muy pormenorizada de lo que permite y lo que no. De esta forma, quien encuentra el documento en la *Web* y desea utilizarlo sabe exactamente lo que el autor le permite hacer y lo que no, sin necesidad de contactos o permisos adicionales. Esta “licencia” queda adosada al instante a su documento electrónico de tres maneras: legible al público en general, en redacción legal y en formato de máquina legible por computadora por un *navegador*. Esta metodología ha sido adoptada ya en veinte países y está en vías de estudio en otros tantos más, entre ellos México, ya que simplifica enormemente el otorgamiento y retención de derechos en publicaciones electrónicas, sin sustituir las leyes de derechos de autor ni el dominio público, pero facilita enormemente la distribución legal y organizada de documentos electrónicos. Ha probado ser ya un excelente complemento para las disposiciones legales en el mundo de las publicaciones electrónicas.

Con el estudio de estas instancias, tanto a nivel de planeación estratégica como al nivel práctico individual, así como su análisis posterior, se tienen los elementos necesarios que buscaba explorar esta obra en materia de derechos de propiedad, tema que como ha podido apreciarse es insoslayable y de capital importancia en el acopio y distribución de información digital.

Capítulo 6

El bibliotecario profesional digital

*“Los bibliotecarios no imaginan
siquiera la cantidad de almas
que han salvado”.*

Barbara Kingsolver.

6.1 PERFIL DEL BIBLIOTECARIO PARA LA BIBLIOTECA DIGITAL

Las bibliotecas digitales requieren de bibliotecarios digitales. Las colecciones digitales deben ser planeadas, seleccionadas, adquiridas, organizadas, distribuidas y preservadas. Los servicios digitales deben ser diseñados, contruidos, operados y mantenidos. Las colecciones digitales y las “Tecnologías de Información y Comunicaciones” son sin duda, herramientas fundamentales para construir estas bibliotecas, pero el personal profesional es elemento indispensable para que todo funcione armónicamente.

Si bien en términos generales los requerimientos y principios para crear y operar una biblioteca digital son básicamente los mismos que para una biblioteca tradicional, las semejanzas terminan allí. Organizar un conjunto de colecciones y servicios digitales representa una serie de acciones y habilidades cotidianas que van más allá de las requeridas para hacerlo con sus correspondientes tradicionales. Los bibliotecarios profesionales que dirigen bibliotecas digitales deben tomar decisiones y realizar acciones para las cuales probablemente no fueron entrenados a fondo durante su estudio profesional. Si a esto sumamos el rápido cambio en los aspectos tecnológicos resulta que lo que se aprendió recientemente queda rebasado en poco tiempo.

De aquí surgen las siguientes preguntas: ¿cuál es el perfil requerido para un profesional de la bibliotecología que trabaja en este ambiente? ¿cuáles características deben permanecer y cuáles deben cambiar? ¿cuál es la capacitación requerida?

Para empezar a contestarlas, es necesario primero partir de un marco de referencia general, y debemos definir qué es un perfil. María

Clemencia Molina y Marta Pérez⁷⁶ afirman que el perfil es: “saber ser y saber hacer puestos de manifiesto en actitudes, conductas o comportamientos en situaciones en las cuales pueda encontrarse un profesional”. Según estas autoras existen varios tipos de perfiles: perfil ocupacional, perfil de competencia, perfil de personalidad, perfil teórico o perfil ideal y perfil profesional.

El perfil profesional del bibliotecario debe involucrar no sólo el perfil ocupacional sino también expectativas de desempeño acordes con las necesidades del medio y una visión acertada de los cambios tecnológicos y sociales que estamos viviendo. En los estudios relativos a perfiles profesionales hay varias maneras de dividir los perfiles; una de ellas ha sido dividir en dos grandes vertientes la caracterización del personal, estableciendo por un lado un enfoque hacia los atributos personales, es decir, las características innatas de un profesional y, por el otro, los atributos profesionales, o sea las características aprendidas durante su formación. Existen por supuesto variadas subdivisiones o agrupamientos de las características según el enfoque de quienes han realizado estos estudios.

De acuerdo con Virginia Cano,⁷⁷ todos esos atributos giran alrededor de tres funciones básicas del bibliotecario profesional, las cuales provienen de la bibliotecología “tradicional” y se conservan en la bibliotecología “digital”:

- El bibliotecario es un guardián del acervo cultural y una de sus funciones primordiales es la de adquirir y preservar el patrimonio bibliográfico.
- El bibliotecario es un creador de sistemas de almacenamiento y una de sus tareas primordiales es proponer la creación de

76 Molina, M^a Clemencia y Pérez, Marta. 1988. “El Currículo y los Perfiles del Profesional en Bibliotecología y Ciencias de la Información”. En: *Revista Interamericana de Bibliotecología*. Medellín: Colombia. 11(1).

77 Cano, Virginia. 1998. “De Bibliotecario a Gestor de Información : ¿Cambio de Nombre o Nuevas Competencias? III encuentro de Directores y II de Docentes de Escuelas de Bibliotecología y Ciencias de la Información del Mercosur, Octubre 29 al 31, 1998. Disponible noviembre 3, 2005 en: <http://www.utem.cl/deptogestinfo/extension.htm>

sistemas de ordenamiento para el patrimonio bibliográfico bajo su custodia.

- El bibliotecario es un mediador entre los sistemas de ordenamiento y los usuarios finales del patrimonio bibliográfico, y uno de sus propósitos principales es facilitar el acceso a la información contenida en el acervo bibliotecario.

Estas tres funciones básicas se conservan en la bibliotecología “digital”, si bien como ya mencionamos, sus formas, características, medios, etcétera, han evolucionado mucho. En el ambiente profesional norteamericano, tan aficionado a los nombres rimbombantes de puestos, funciones y títulos, así como de acrónimos, se ha desatado desde hace algunas décadas, y con notorio énfasis en la última, una notable tendencia a *bautizar* con un sinnúmero de neologismos a las “nuevas” funciones o a los “nuevos” bibliotecarios, en un afán de destacar la novedad y especialidad de las funciones, las habilidades o las actitudes. Linda Braun⁷⁸ afirma

“para poder satisfacer las necesidades del cliente en la actualidad, la mayoría del personal de la biblioteca debe estar entrenado para proveer materiales tanto en formatos tradicionales como digitales. Un sinnúmero de bibliotecarios han integrado la tecnología a su trabajo sin ningún cambio en la descripción de su puesto o título. Los bibliotecarios de consulta usan la Internet, catálogos y bases de datos en-línea para ayudar a los usuarios a localizar materiales. De todos modos, los bibliotecarios ya no son denominados así. El incremento en sus responsabilidades y títulos proporciona un marco laboral para los papeles tecnológicos que ellos desempeñan dentro de la biblioteca. Las revistas profesionales y las listas electrónicas de discusión reflejan esos nuevos papeles”.

78 Braun, Linda W. 2002. “A Librarian by Any Name”. En: *Library Journal*. 127 (2):46. Disponible noviembre 3, 2005 en: <http://www.libraryjournal.com/index.asp?layout=articleArchive&articleId=CA191647&display=searchResults&stt=001>

Además, esta costumbre ha permeado a Europa y América Latina, y se ha creado una Babel de descripciones y nombres para las personas que dedican su vida profesional a la muy noble y milenaria tarea de realizar las tres funciones enunciadas previamente, simplemente para hacer accesible la información al público. Así, hemos visto nacer al *documentalista*, al *bibliotecario de medios*, al *especialista en información*, al *gerente de información*, al *cibertecario*, al *gestor de información* y al *del conocimiento*, al *científico de la información*, al *administrador del conocimiento*, al *profesional de búsquedas en-línea*, al *especialista en acceso*, etcétera. De hecho, hay un sitio en Internet que consigna más de 300 nombres para las personas que se dedican como profesionales a la información y la documentación. Todos los nombres son reales y han sido extraídos de descripciones de puestos genuinos de la vida real.⁷⁹ Ana María Martínez Tamayo parodia este hecho maravillosamente en su *Babel Bibliotecaria*.⁸⁰

Aquí se observan dos tendencias: una que tiende a denominar con nuevos nombres a la función bibliotecaria profesional, para distinguir esas nuevas características como única manera de establecer una diferencia entre el bibliotecario “tradicional” y el nuevo bibliotecario que tiene una serie de aptitudes y habilidades tecnológicas, y otra que insiste en que la profesión sigue siendo la bibliotecología y su profesional, el bibliotecario, incluyendo por supuesto las habilidades y aptitudes mencionadas.

Yo en lo personal me inclino a este segundo enfoque: no considero necesario ni recomendable renunciar a una denominación que ya hunde sus raíces en milenios de tradición, sólo por el hecho de querer dar una nueva imagen del perfil profesional bibliotecario.

Los médicos deciden seguir llamándose médicos. Nadie pone en duda los vertiginosos avances de la medicina en el siglo xx, en particular hacia fines del mismo. Los médicos de hoy en día no curan ya con las técnicas de la época de Hipócrates, no hacen sangrías ni aplican

79 Job Title Generator for library and Information Sciences Professionals. Disponible noviembre 3, 2005 en: <http://lamar.colostate.edu/~mach/jobtitles/>

80 Martínez Tamayo, Ana María. *Babel Bibliotecaria*. Disponible noviembre 3, 2005 en: <http://www.fahce.unlp.edu.ar/departamentos/dhubi/paginas/clasificacion/Babel%20bibliotecaria.htm>

ventosas, ni curan como lo hacían hace treinta o cincuenta años. Es obvio que usan para su noble tarea los adelantos más notables de la tecnología y del conocimiento humano para hacerlo. De hecho, he escuchado recientemente a varios médicos afirmar que lo que ha avanzado en las últimas décadas es la ingeniería médica, no la medicina. Ahora operan con láseres en lugar de bisturíes, deshacen cálculos con litotriptores, exploran el cuerpo con tomógrafos, ciclotrones y medicina nuclear. Complejos medicamentos insospechados hace dos décadas complementan su tarea. La función sigue siendo la misma que hace varios miles de años: preservar la vida y ayudar a recuperar la salud de sus pacientes. Las técnicas y las herramientas ya no lo son. No obstante, yo no escucho a los médicos denominarse *científicos de la salud*, ni *cibercirujanos*, ni *gestores de la salud*, ni *administradores de la sanidad*, ni nada por el estilo. Siguen llamándose “médicos”.

Eso sí, se han especializado y preparado concienzudamente. Ostentan con orgullo siempre una especialidad o un grado además de su título de médico cirujano. Se actualizan permanentemente; aprenden a usar nuevas herramientas y dominan nuevos aparatos y materiales que saben aplicar a los requerimientos de sus pacientes. Aprenden nuevas técnicas y conocen nuevos medicamentos. Nadie espera que curen sólo con estetoscopio, rayos X, cloroformo y penicilina y ellos lo saben.

La ingeniería, disciplina mucho más reciente, nació con la revolución industrial. Toma su nombre de los “ingenios” —*engines*— de esa época, nombre que se usó para denominar cualquier maquinaria compleja que realizara una tarea dada, como las locomotoras o los tornos. Todavía hoy se sigue llamando en español “ingenio” a un complejo de molienda de caña, en extensión a su complicada maquinaria de origen ancestral, y *engineer* en inglés a un maquinista de locomotora.

Los ingenieros, es decir, los profesionales de la ingeniería, extendieron su campo, entre otros, a la construcción con la ingeniería civil, a la mecánica, la eléctrica, la petrolera, la electrónica y recientemente, a la ingeniería en computación. Construyen desde complicadas presas y rascacielos, plataformas petroleras, equipos médicos, naves espaciales y satélites, hasta aviones y computadoras. Por supuesto, con mucho mayores conocimientos y un impresionante abanico de máquinas, herramientas y dispositivos con lo más moderno de la tecnología, incorporan

siempre lo mejor y más reciente de ello a su quehacer. No usan más la regla de cálculo y la tabla de logaritmos. No construyen ya nada con la tecnología de hace veinte años, no digamos con la de hace cincuenta... pero siguen llamándose orgullosamente “ingenieros”, a pesar de que el concepto original de “ingenio” o maquinaria sea más que obsoleto. No buscan ser denominados *científicos de la ingeniería*, ni *administradores tecnológicos*, ni *gestores de la técnica*, ni *ciberingenieros*, ni *especialistas en la tecnociencia de la ingeniería*. De hecho, los ingenieros son de los pocos profesionales que aun ostentando una maestría, doctorado o especialización no se molestan si alguien los llama “ingeniero”; tanto así valoran su profesión, la cual sin duda tiene los mismos requerimientos de actualización continua, de capacitación, de incorporación de nuevas tecnologías y herramientas a su quehacer, como tantas otras de las profesiones actuales.

Como estos dos casos, podríamos enunciar muchos otros, basten como ejemplo para sacar una enseñanza de ellos. No creo que el bibliotecario tenga que abandonar su secular denominación en aras de una imagen nueva. ¿necesita tener una especialidad, si no un posgrado? sin duda. ¿requiere de nuevas aptitudes y actitudes en su profesión? es obvio que sí. ¿Debe actualizarse continuamente a lo largo de su vida profesional? huelga decirlo. ¿Debe dominar nuevas herramientas e incorporarlas a su quehacer cotidiano? ni hablar.

De esta forma podemos presentarnos como un bibliotecario, pero podemos y debemos agregar que tenemos una o varias especialidades. Debemos poder decir que tenemos buenas aptitudes y habilidades tecnológicas. Debemos poder afirmar que sabemos diseñar y administrar nuevos y modernos servicios de información en variados formatos y, en suma, que somos capaces de concebir, construir y operar una biblioteca digital, en toda la cabal extensión del término. Y que podemos hacerlo mejor que cualquier otro “profesional” de la información.

David Stoker⁸¹ resume esto diciendo: “hoy en día, a pesar de los cambios en los currículos que han ofrecido los departamentos de edu-

81 Stoker, David. 1999. “Wanted? An Innovative and Visionary Evidence Based/ Knowledge Management Librarian.” En: *Journal of Librarianship and Information Science*. 31(2):67-69.

cación en información durante las tres últimas décadas existe un sorprendente y considerable número de puestos que son designados como ‘bibliotecario’ de un tipo o de otro...”

Al margen de las preferencias de denominación de cada quien, para fines de este trabajo he decidido, por lo anteriormente expuesto, que debemos definir el perfil y la caracterización de un profesional que denomino *Bibliotecario Digital*. “Bibliotecario” porque sigue manteniendo la esencia de la profesión, y “digital” para enfatizar este enfoque deseado.

Con esta reflexión previa comienzo el análisis detallado de ese perfil deseado sobre actual profesional de la bibliotecología que concibe y opera las bibliotecas digitales.

6.2 EL BIBLIOTECARIO DIGITAL EN NORTEAMÉRICA

La caracterización más conocida para este profesional bibliotecario es la creada por la “Asociación de Bibliotecas Especializadas” —*Special Libraries Association* o simplemente SLA— de la unión americana. Si bien es un estudio realizado en los Estados Unidos y enfocado a bibliotecas especializadas es sin duda uno de los hitos más conocidos para comenzar un análisis acerca de este tema [Special Libraries Association, 1996]. He aquí un resumen de la caracterización:

Competencias personales

El bibliotecario especializado:

- 1) Está comprometido con la excelencia en el servicio. Busca la retroalimentación de su trabajo; realiza estudios de usuarios y trata continuamente de averiguar si la percepción acerca de la información que el y sus compañeros encuentran es valiosa y relevante. Se enorgullece de un trabajo bien hecho, y aprovecha su experiencia profesional para pensar en mejorar los servicios.

- 2) Busca nuevos retos y ve nuevas oportunidades dentro y fuera de la biblioteca. Acepta jugar nuevos papeles en la organización que requieran un líder de la información. Usa su conocimiento y habilidades para resolver los problemas de información en una amplia variedad de situaciones. Expande la colección más allá de los libros y revistas. Crea nuevos servicios.
- 3) Tiene visión de conjunto. Está consciente de que la información es parte de un proceso creativo para las personas y las organizaciones. Ve a la biblioteca y a sus servicios de información como parte de un gran proceso de ayuda para la toma de decisiones. Le da la más alta prioridad a las demandas urgentes de información que son críticas para la competitividad de las organizaciones. Está al tanto de tendencias mundiales en cuanto a información para negocios o proyectos, y realinea a su biblioteca proactivamente para contender con esas nuevas demandas.
- 4) Negocia asociaciones y alianzas. Se asocia con otras bibliotecas y sistemas de información para potenciar proyectos, servicios y colecciones. Busca alianzas con proveedores de información para mejorar sus colecciones y servicios. Logra alianzas con entidades académicas de investigación para realizar proyectos vanguardistas en materia de información.
- 5) Crea un ambiente de confianza y respeto mutuo. Conoce las fortalezas y habilidades de su personal y las impulsa y desarrolla. Trata a todos con respeto y actitud profesional. Realiza su tarea en tiempo y forma e impulsa a otros a hacer lo mismo. Crea un ambiente enfocado a la obtención de resultados en los que las contribuciones de todos son valoradas y reconocidas.
- 6) Posee habilidades de comunicación efectivas. Sabe escuchar y asesora a otros sobre cómo lograr sus propias soluciones a los problemas. Apoya y comparte la planeación participativa. Encabeza reuniones productivas. Presenta sus ideas con claridad y entusiasmo. Sabe escribir documentos útiles para la organización. Usa un lenguaje sencillo.
- 7) Sabe trabajar en equipo. Está consciente de la fuerza que tiene un equipo. Asume su parte en tareas conjuntas. Evita un ambiente de trabajo basado en genialidades individuales. Asesora a otros

miembros del equipo y pide asesoría cuando la requiere. Desalienta la competencia estéril.

- 8) Posee liderazgo. Aprende y cultiva las cualidades de un buen líder y sabe cuándo ejercer el liderazgo y cuándo delegarlo. Assume ese liderazgo en eventos dentro y fuera de la biblioteca. Apoya e impulsa la creación de nuevos líderes que lo sucedan.
- 9) Planea, prioriza y se enfoca en los aspectos críticos. Sabe que la planeación es vital para optimizar el uso de los recursos que se tienen a la mano. Sabe ajustar los planes para lograr resultados. Revisa sus metas y prioridades regularmente para retroalimentar sus planes.
- 10) Está comprometido con el aprendizaje continuo y planea su carrera. Está consciente de que su carrera profesional depende en gran parte de un proceso de aprendizaje y actualización a lo largo de su vida. Busca siempre oportunidades para nuevas capacitaciones y el desarrollo de nuevas habilidades. Quiere especializarse en varias facetas de su carrera y va logrando con ello su autoestima profesional.
- 11) Tiene habilidades empresariales y crea nuevas oportunidades. Reconoce que en un mundo laboral cambiante, la capacidad empresarial y la habilidad para desenvolverse en grandes y pequeñas empresas son atributos esenciales. Acepta retos laborales muy variados, como contratos y proyectos, además de trabajos de tiempo completo. Aprovecha ese espíritu empresarial para innovar en relación con bienes y servicios en su biblioteca.
- 12) Reconoce el valor de la asociación profesional y la solidaridad. Participa en asociaciones gremiales y profesionales. Comparte sus conocimientos y habilidades. Valora la necesidad de participar en foros para intercambiar ideas entre profesionales y expresa su opinión en la implementación de programas y políticas que benefician a la bibliotecología.
- 13) Es flexible y positivo en esta época de cambio. Se atreve a afrontar nuevas responsabilidades y responde a las necesidades de cambio. Mantiene una actitud abierta y positiva, y ayuda a otros a hacerlo. Nunca dice que algo es imposible; busca soluciones. Ayuda a otros a desarrollar sus innovaciones, y siempre está atento

a nuevas formas de hacer las cosas. Aprovecha la tecnología como activador que es de nuevas ideas, productos y servicios.

Competencias profesionales

El bibliotecario especializado:

- 1) Tiene conocimiento y experiencia de los recursos de información, y tiene la habilidad de evaluarlos críticamente así como de filtrarlos. Conoce los mejores textos, revistas y recursos electrónicos en las principales áreas. Evalúa y selecciona recursos claves para su biblioteca. Controla la sobreoferta de información seleccionando lo más útil y relevante para sus usuarios.
- 2) Posee un conocimiento profundo de la especialidad de la organización para la que trabaja. Se va adentrando en la materia de especialidad de la biblioteca para la que trabaja e incluso cursos al efecto. Monitorea el área de su biblioteca leyendo las revistas de mayor trascendencia, y ello le permite ir diseñando servicios cada vez más especializados y responder mejor a las necesidades de sus usuarios.
- 3) Desarrolla y administra servicios de información accesibles, adecuados y rentables alineados con las estrategias de su organización. Desarrolla un plan estratégico de información acorde con el plan de su institución. Establece la administración, supervisión y presupuesto eficazmente. Forma un adecuado equipo humano de información. Coordina búsquedas complejas a lo largo de variadas fuentes. Va construyendo una colección núcleo, tesauros e índices especializados.
- 4) Proporciona excelente entrenamiento y apoyo para los usuarios de la biblioteca y sus servicios. Les enseña a acceder a los recursos en red a otros empleados. Desarrolla cursos especializados en el tipo de información requerida por la biblioteca de su especialidad. Resuelve problemas de usuarios y otros empleados que acceden a servicios de información. Provee referencia y asistencia en-línea.

- 5) Asesora acerca de requerimientos de información; diseña y opera productos y servicios de información con valor agregado para necesidades específicas. Elabora encuestas de usuarios y mercado para presentar soluciones a las necesidades detectadas. Identifica necesidades de información al formar parte de equipos multidisciplinarios.
- 6) Usa la apropiada tecnología de información para adquirir, organizar y diseminar información. Crea catálogos en línea. Liga los catálogos a servicios de entrega documental. Trabaja junto con los equipos informáticos en ayudar a establecer los equipos y programas adecuados para acceder a los recursos informáticos. Se mantiene al día acerca de los avances de tecnologías de información y las maneras de entregársela al usuario.
- 7) Usa enfoques administrativos y gerenciales adecuados para sensibilizar a los mandos superiores acerca de la importancia de los servicios de información. Desarrolla modelos de negocio para distintos servicios en la biblioteca. Busca mejorar la calidad de los servicios. Demuestra que los servicios documentales agregan valor a la organización.
- 8) Desarrolla productos de información especializados para uso tanto de su organización como de otros individuos. Crea bases de datos de documentos propios de su área, tales como reportes, artículos, manuales, etcétera. Crea archivos documentales en texto completo. Concibe una buena página *web* documental para su organización, con buenas ligas a otros sitios pertinentes. Participa en proyectos que exploten el capital intelectual y documental de la organización.
- 9) Evalúa el efecto del uso de la información y realiza investigaciones para solucionar problemas sobre el manejo de la información. Diseña programas de evaluación informativa, frecuencia de uso de los servicios, satisfacción de usuarios, impacto de la información.
- 10) Mejora continuamente los servicios de información en respuesta al cambio de las necesidades. Monitorea los estándares de la industria y los disemina entre las personas clave en la organización. Afina los servicios de información de acuerdo

con el cambio en las necesidades. Entrega los documentos en plazos cortos. Monitorea las adquisiciones de productos de información para asegurar su rentabilidad.

- 11) Es un miembro efectivo del grupo gerencial de la organización y un consultor efectivo de la misma en materia de información. Participa en planes estratégicos de su organización. Vigila el adecuado movimiento dentro de derechos de propiedad. Negocia contratos con vendedores de productos y servicios de información. Desarrolla políticas informativas dentro de su organización.

Nadie puede negar que lo que ha sido así descrito es el perfil de alguien con grandes aptitudes y actitudes tecnológicas y empresariales. Y tampoco que cuenta puede dudarse de que la persona aquí descrita es un bibliotecario profundamente profesional.

6.3 EL BIBLIOTECARIO DIGITAL EN EUROPA

Continuemos ahora con un resumen de lo expresado en la Unión Europea alrededor del perfil deseado de los bibliotecarios “digitales” para los nuevos proyectos que se desarrollan en esa región del mundo. La Comunidad solicitó a su departamento responsable de educación y entrenamiento, conocido como la *Dirección General XXII*, la elaboración de una *Relación de Eurocompetencias en Información y Documentación*, producto del proyecto DECIDOC, en el que participaron nueve asociaciones profesionales europeas para promover la formación profesional [European Council of Information Associations, 1999] y que cristalizó a través de dos proyectos de desarrollo, denominados *Sócrates* y *Leonardo da Vinci*. Este último generó un documento que es el más extenso y aceptado a nivel europeo de lo que debe normar las *eurocompetencias*, como ellos las llaman, de los profesionales de la información en los próximos años.

De acuerdo con este documento, y a manera de resumen, éstas serán las principales tareas de los profesionales de la información en el futuro cercano:

- Gestionar las Bibliotecas Digitales bajo el concepto de centros de recursos de acceso para el aprendizaje y la investigación en este nuevo entorno tecnológico.
- Organizar el conocimiento y la información digital.
- Difundir la información procedente de recursos digitales.
- Proporcionar servicios digitales/electrónicos de información y referencia.
- Identificar y documentar nuevas tendencias en el acceso a la documentación digital y sus tecnologías.
- Manejar las tareas de digitalización masiva, el proceso de almacenamiento digital y la preservación o conservación digital.
- Garantizar un acceso y recuperación global del conocimiento digital.
- Catalogar y clasificar los documentos y el conocimiento digital.

La aparición de estas nuevas funciones hace necesaria la definición de nuevos perfiles profesionales, con la consiguiente descripción de competencias, entendiendo como tales: “el conjunto de capacidades necesarias para el ejercicio de una actividad profesional y el dominio de los comportamientos correspondientes”. Sus componentes son conocimientos teóricos, prácticos y aptitudes, o “el conjunto de conocimientos, destrezas y aptitudes cuya finalidad es la realización de actividades definidas y vinculadas a una determinada profesión”.

Definición de competencias y niveles.

Este documento define cuatro grandes grupos de competencias:

- Conocimientos específicos de la información y documentación.
- Conocimientos sobre comunicación.
- Competencias sobre gestión y organización.
- Otros conocimientos aplicados a la información y documentación.

En estos cuatro grupos se detallan treinta campos de competencia, escalonados, a su vez, en cuatro niveles significativos:

- NIVEL1 Sensibilización** , el profesional conoce la existencia de los elementos que forman una actividad o función concreta. Es necesaria una cultura general básica en el campo de competencia: conocimiento del vocabulario básico, capacidad para ejecutar algunos elementos prácticos o concretos.
- NIVEL2 Conocimientos de las prácticas** . El profesional dispone de instrumentos de lectura y de expresión de los fenómenos estudiados. Es el primer nivel profesional. Sabe manejar herramientas básicas, efectuar trabajos de ejecución especializados o repetitivos y transmitir propuestas prácticas.
- NIVEL3 Dominio de herramientas** . El profesional conoce la existencia de técnicas, sabe definir las, hablar de ellas y domina su utilización. Es capaz de interpretar una situación y formular un juicio que implique cambios o modificaciones, y está capacitado para crear una herramienta.
- NIVEL4 Dominio metodológico** . El profesional utiliza una técnica determinada, pero puede aplicarla en circunstancias distintas, adaptarla a otras tareas, encontrar otros campos de aplicación, concebir mejoras u otras técnicas más elaboradas o mejor adaptadas. Puede concebir herramientas o productos nuevos y adoptar un enfoque estratégico o general de su actividad. Percibe la complejidad de las situaciones y puede encontrar soluciones originales adaptadas a ellas.

Todos los grupos de competencias son aplicables al perfil de bibliotecario digital. No obstante, nos hemos centrado en aquellas competencias que tienen relación directa con el perfil concreto que tiene que desarrollar el bibliotecario en un escenario tecnológico/informacional.

Detalle de las competencias en tecnologías de la información.

Nivel 1

- Utilizar un equipo, una herramienta informática o una aplicación de oficina automatizada normal.
- Utilizar las funcionalidades principales de un sistema de explotación documental.
- Comprender un vocabulario básico de tecnologías de la información.
- Poder acceder y consultar un sistema de información.
- Detectar y comprender anomalías simples en el funcionamiento de un sistema informático.
- Manejar básicamente los sistemas de búsqueda de información.

Nivel 2

- Tener conocimientos avanzados sobre aplicaciones de oficina automatizada.
- Diagnosticar la naturaleza de una avería básica del equipo.
- Contar con la capacidad para definir las características básicas de primer nivel del *hardware* / *software*.
- Poder instalar un puesto; esto es, equipo informático; un programa informático o una aplicación de oficina automática corriente en un equipo.
- Poder hacer el mantenimiento básico de un parque informático constituido por equipos homogéneos bajo la responsabilidad de un administrador.
- Diferenciar y utilizar los *softwares* especializados de gestión, búsqueda documental, conversión, compresión de archivos, digitalización y edición de documentos.
- Tener capacidad para apoyar el desarrollo de aplicaciones documentales.
- Garantizar el seguimiento de la explotación y el mantenimiento de los sistemas.
- Poder elaborar y actualizar páginas *Web*.
- Tener conocimientos básicos para digitalizar documentos.

Nivel 3

- Poder hacer la instalación y parametrización de *software cliente* —local— para acceder a sistemas de información comunes en un parque informático heterogéneo.
- Saber redactar un pliego de especificaciones funcionales para una aplicación documental informatizada o un sistema de gestión de información.
- Desarrollar aplicaciones documentales informatizadas cualesquiera que sean los medios de acuerdo con la política general de la institución en esta materia.
- Poder utilizar un lenguaje de programación o “macrocomandos” para escribir programas elementales que permitan extender determinadas funciones.
- Gestionar los derechos de acceso de los usuarios garantizando su confidencialidad y privacidad.
- Diseñar de actividades formativas encaminadas a la instrucción de los usuarios en las tecnologías de la información.
- Tener capacidad para crear o mantener un sitio *Web* o una *Intranet*.
- Saber elaboración de productos digitales de calidad profesional.

Nivel 4

- Elaborar y proponer un plan tecnológico que garantice la viabilidad de los recursos electrónicos / digitales de la biblioteca.
- Coordinar la política tecnológica de la biblioteca con la política tecnológica de la institución.
- Establecer acuerdos y convenios que faciliten el intercambio tecnológico entre diferentes instituciones.
- Buscar fuentes y fórmulas de financiamiento externas. Proponer políticas encaminadas a facilitar el acceso a todos los recursos tecnológicos disponibles.
- Diseñar campañas para la formación de bibliotecarios y usuarios en tecnologías de la información.
- Proponer políticas que aseguren la conservación del patrimonio bibliográfico documental.

Las treinta competencias comunes y relacionadas que pueden ser ejercidas tanto por el “bibliotecario digital” como por otros perfiles de bibliotecario, y que se encuentran consignadas en el documento de euro-competencias son las siguientes:

1. Conocimientos específicos de información y documentación:

- Interacción con los usuarios.
- Conocimiento del entorno profesional de la información y documentación.
- Conocimiento del marco jurídico y administrativo europeo de la gestión de la información.
- Identificación y validación de fuentes de información.
- Gestión de colecciones y fondos.
- Tratamiento físico de documentos.
- Análisis y representación de la información.
- Organización y almacenamiento de la información.
- Búsqueda de información.
- Elaboración y Difusión de la Información.

2. Conocimientos sobre comunicación:

- Comunicación oral.
- Comunicación escrita.
- Dominio de una lengua extranjera.
- Comunicación visual y sonora.
- Comunicación interpersonal.
- Comunicación institucional.
- Tecnologías de información: informática.
- Tecnologías de información: telecomunicaciones.
- Técnicas de producción y edición.

3. Competencias relevantes sobre gestión y organización:

- Técnicas de gestión administrativa.
- Técnicas comerciales.
- Técnicas de mercadotecnia.
- Técnicas de adquisición.
- Técnicas de gestión micro-económica.

- Técnicas de instalación, acondicionamiento y equipamiento.
- Técnicas de planificación y gestión de proyectos.
- Técnicas de diagnóstico y evaluación.
- Técnicas de gestión de recursos humanos.
- Técnicas de formación.

4 . Otros conocimientos aplicados a la información y documentación:

- Todos los conocimientos útiles para la gestión de la información.

Aptitudes. Además de los conocimientos —teóricos y prácticos—, se precisan una serie de “aptitudes” que deben acompañar a los profesionales y que son necesarias para el adecuado desempeño de los diferentes cometidos. A veces, estas aptitudes son aún más necesarias que los conocimientos teóricos y prácticos para asegurar la eficacia de un profesional y su buena inserción en el medio de trabajo. Estas aptitudes se pueden resumir en:

- Capacidad de adaptación.
- Capacidad de análisis y comunicación.
- Curiosidad intelectual y espíritu crítico.
- Capacidad de decisión.
- Capacidad para la iniciativa y el trabajo en equipo.
- Capacidad de síntesis y de organización.

El documento europeo finaliza diciendo: “Este documento es una herramienta básica para la definición de las competencias requeridas para un ***bibliotecario digital***. Debe ser revisado periódicamente para recoger las innovaciones y cambios que permitan a los profesionales adaptarse a las nuevas tecnologías y su evolución. Puede adaptarse para ayudar a las bibliotecas a establecer sus propias competencias.” Hasta aquí el resumen del documento europeo de “Eurocompetencias”.

6.4 EL BIBLIOTECARIO DIGITAL EN LATINOAMÉRICA

Continuemos ahora con algunas revisiones a las posturas establecidas en América del Sur, quienes en Latinoamérica han discutido y reflexionado más sobre el tema que en México. La Universidad Tecnológica Metropolitana Chilena, que tiene la escuela de bibliotecología en Santiago, establece los siguientes lineamientos, tomados a su vez de los estudios de Benítez de Vendrel de 1999, y que fueron adoptados previamente en Argentina. De acuerdo con ellos la escuela establece en relación con el tema de las competencias, un conjunto de tres grandes capacidades del profesional bibliotecario, que se traducen en: un saber hacer —competencias operacionales—, un saber sobre —competencias cognitivas— y un saber que se relaciona con una conciencia sobre —competencias actitudinales—. *De acuerdo con el mercado laboral*, indistintamente de dónde provenga la oferta de trabajo y de la función que el profesional bibliotecario deba desempeñar en ella, se requiere del bibliotecario que tenga las siguientes competencias:

- **Competencias cognitivas** : Cultura general, manejo del idioma inglés, sólidos conocimientos bibliotecológicos y sobre administración y gestión de biblioteca, *software* para bibliotecas, buen conocimiento en computación y redes de información a nivel de usuario.
- **Competencias operacionales** : El profesional bibliotecario debe: ser buen gestor de bibliotecas, dominar el manejo de *software* para bibliotecas, tener dominio del manejo computacional a nivel de usuario, así como de redes de información, concretamente *Internet*.
- **Competencias actitudinales** , el mercado laboral exige que el profesional tenga capacidad para trabajar en equipos disciplinarios y multidisciplinarios, que muestre flexibilidad al cambio y tenga facilidad para establecer buenas relaciones interpersonales.

De acuerdo a lo establecido *por otros bibliotecarios*, las competencias actualmente exigibles para el profesional de la biblioteca, son:

- **Competencias cognitivas:** El profesional bibliotecario debe poseer sólidos conocimientos bibliotecológicos en materia de información y comunicación, análisis de información, descripción bibliográfica, análisis de sistemas, formatos, administración y gestión de unidades de información, formulación de proyectos de información, fuentes de información, organización de archivos, educación de usuarios, redes de información, tecnologías de información aplicables a las bibliotecas. En materia de conocimientos generales es fundamental una cultura general amplia y buen dominio del inglés.
- **Competencias operacionales :** Debe tener un buen manejo de herramientas computacionales y tecnológicas, y dominio del uso de la red Internet.
- **Competencias actitudinales** que debe poseer el profesional: se destacan la proactividad, la asertividad, la iniciativa y la creatividad. De las capacidades mencionadas, la proactividad, que se define como la capacidad de ser promotor del cambio y no actuar por reacción a él, es la competencia más recurrente. Otras competencias actitudinales comunes son las habilidades comunicacionales, buenas relaciones interpersonales y la capacidad para ponerse en el lugar del usuario. La ética y la vocación son valores que todo profesional debe poseer.

Cognitivas (saber)	Operacionales (saber hacer)	Actitudinales (saber actuar)
<p>Cultura general. Dominio del inglés y conocimiento de otro idioma. Tecnología de la información en general. Bases de datos. Sistemas operativos y herramientas computacionales. Redes de información. Internet. Dispositivos de almacenamiento y recuperación de información. Tecnologías de información para bibliotecas.</p>	<p>Tecnologías de la información aplicables a bibliotecas. Manejo de software para bibliotecas. Capacidad de administración y gestión de bibliotecas. Tecnologías de información en general. Manejo de redes de información. Manejo de computación nivel de usuarios. Redes de información: Internet. Dispositivos de almacenamiento y</p>	<p>Dinamismo. Entusiasmo. Espíritu creativo. Innovación e iniciativa. Diplomacia y buenas relaciones interpersonales. Capacidad de comunicación y trato. Trabajo en equipo. Responsabilidad. Valores sólidos. Trabajo bajo presión. Vocación de servicio. Liderazgo.</p>

Cognitivas (saber)	Operacionales (saber hacer)	Actitudinales (saber actuar)
Software para bibliotecas. Servicios de alerta. Catalogación digital.	recuperación de información.	Orientación al logro. Flexibilidad al cambio. Respeto por los demás y por la profesión. Conducta ética. Responsabilidad. Compromiso.

FIGURA 6.1 – TABLA DE COMPETENCIAS PROFESIONALES DEL BIBLIOTECARIO DIGITAL.

Cecilia Braslavsky sostiene que existen competencias aglutinantes que corresponden a diversos grupos de competencias y se estructuran en torno a un elemento central. Desde este punto de vista, se pueden identificar los siguientes grandes grupos: intelectuales, prácticas, interactivas y sociales, éticas y estéticas. Creo apropiada la inclusión de una breve referencia a ellas, a fin de clarificarlas para evitar confusiones y arrojar luz sobre los procesos que se requieren en cada especialidad.

De acuerdo con esta autora las *competencias intelectuales* hacen referencia a los procesos cognitivos internos necesarios para simbolizar, representar ideas, imágenes, conceptos u otras abstracciones; también suelen aparecer bajo el rótulo de competencias analíticas, creativas y metacognitivas. Las *competencias prácticas* se refieren a un “saber hacer”, es decir a poner algo en acción. Implican en cierta medida nociones de administración espacial, temporal y de recursos. Las *competencias interactivas* implican en sí mismas la capacidad de los sujetos de participar como miembros de grupos de referencia —la familia, grupos de pares—. Y suponen un entrenamiento en el consenso, en el ejercicio y la aceptación del liderazgo; es decir, en la capacidad de enseñar y de aprender con otros. Se la suele llamar *competencia social* —aunque otros teóricos le suelen dar un concepto más abarcativo que incluye lo macrosocial—. Las *competencias éticas* son las rectoras para discernir lo bueno de lo malo en el amplio y complejo espacio en el que pugnan los órdenes sociales; por ejemplo el derecho a la vida, las pautas culturales, las creencias, las religiones, el amor, la educación. Por último, hacemos referencia a las *competencias estéticas* que aluden

a la capacidad de distinguir lo que en el plano de los valores es bello o no; por supuesto que éstas implican también los condicionamientos sociales y culturales.

A continuación, presento la siguiente agrupación, según José Augusto Guimarães, del perfil que debe tener el profesional bibliotecario de acuerdo con la visión brasileña del mismo. [Guimarães, 1998]:

El bibliotecario profesional moderno debe tener:

- Flexibilidad.
- Capacidad de análisis.
- Creatividad.
- Liderazgo.
- Dinamismo.
- Responsabilidad.
- Visión y actuación inter y multidisciplinaria.
- Profesionalismo.
- Ética.
- Dominio de una técnica de organización moderna del conocimiento.
- Visión política en el área de información.
- Uso de la información para ventaja competitiva de las organizaciones.
- Uso de la información para el desarrollo social y humano.
- Entrenamiento en recursos informacionales e informáticos.
- Espíritu y acción investigadora.
- Compromiso con la creación de nuevos mercados de trabajo
- Objetividad y actitud crítica; claridad y precisión.
- Agilidad mental.
- Motivación interna para disfrutar del trabajo como satisfacción personal.
- Habilidad para la solución de problemas.
- Valor para enfrentar los riesgos, sin temor a la competencia.
- Habilidad en tecnologías de información como herramienta profesional.
- Visión gerencial en el área de información.
- Habilidad en el acceso de diversos soportes y formatos de información.
- Aptitud y Actitud comunicativa y educativa.
- Compromiso con la calidad de los servicios informativos.
- Habilidad de adelantarse a las innovaciones.
- Actualización constante.
- Capacidad para ampliar los horizontes de su organización.
- Habilidad para aprender de sus usuarios.
- Adaptabilidad a su ambiente.
- Empeño en aprovechar las oportunidades en favor de la sociedad.
- Variados profesionales integrados a su equipo de trabajo.
- Capacidad de agregar valor a la información.
- Amor por su profesión.

FIGURA 6.2 – TABLA DE APTITUDES DEL BIBLIOTECARIO PROFESIONAL DIGITAL. [GUIMARÃES, 1998]

Como puede observarse esta visión brasileña no difiere mucho de las anteriores en cuanto a cuáles deben ser las características que definen a un profesional moderno de la bibliotecología.

Me parece que con estas distintas visiones y caracterizaciones de lo que se espera de un bibliotecario profesional en la actualidad es suficiente para entender la responsabilidad del perfil que los bibliotecarios *digitales* deben tener frente a la sociedad; en particular ante las organizaciones que los contratan y ante los usuarios a quienes atienden.

Por supuesto que esto implica una reflexión conjunta entre los profesionales de la bibliotecología actuales en el campo del ejercicio profesional y de quienes tienen a su cargo la formación curricular de los futuros bibliotecólogos en las escuelas correspondientes. Ello implica que los planes de estudio, balance de cargas curriculares académicas, etcétera, tengan que ser replanteados en algún momento y cada vez más con mayor frecuencia para poder responder a estos perfiles profesionales.

Roberto Garduño⁸² lo resume así: "También se advierte que el constante desarrollo de la ciencia y la tecnología en las diversas áreas profesionales ha propiciado una acelerada generación del conocimiento, situación que dificulta su estudio permanente en niveles óptimos de aprendizaje. En este sentido se percibe que los fenómenos surgidos con las tecnologías y los requerimientos del mercado laboral han dado lugar a la formación de profesionales de alto nivel. Al mismo tiempo se ha generado el requerimiento de incidir en una mejor preparación de los recursos humanos, lo cual es propiciado por la orientación hacia lo global, a democratizar la educación, a la constante generación de recursos documentales y a incrementar las posibilidades que otorgan los desarrollos tecnológicos [...]"

82 Garduño, Roberto. 2004. "La Educación a Distancia como elemento del Desarrollo de la Sociedad del Conocimiento". En: *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, vol. 37, no. 18. Diciembre 2004. pp. 7-8. ISSN: 0187-358-X. Disponible noviembre 3, 2005 en: <http://www.ejournal.unam.mx/iibiblio/vol18-37/IBI03701.pdf>

...Frente a los nuevos desafíos la bibliotecología requiere diversificar su campo de conocimiento y desarrollar nuevas alternativas de formación tomando en cuenta los avances del conocimiento en el área”.

Finalmente, concluyo este capítulo con esta reflexión de Jesse Shera⁸³ acerca del estado del arte de la profesión bibliotecaria, que a pesar de tener ya quince años me parece que se encuentra cada día más vigente dentro del contexto actual:

“De todas las profesiones la del bibliotecario es seguramente la más derivativa y sintética, dependiente sobremanera de las disciplinas más formales para la derivación de su propia estructura teórica y cuerpo de práctica [...] esta calidad le ha dado a la bibliotecología una posición estratégica única de liderazgo en la integración del conocimiento humano y podría hacer de la bibliotecología una gran fuerza unificadora, no solo en el mundo del saber sino en el de toda la vida humana”.

83 Shera, Jesse H. 1990. “Lo que el Bibliotecario Necesita Saber”. En: *Los Fundamentos de la Educación Bibliotecológica*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas. pp. 201-231.

"Afortunadamente, la mente humana no es del todo elástica; cada vez que es expandida por una nueva idea, nunca regresa a sus dimensiones anteriores".

RESUMEN

Como fue enunciado desde el prólogo de este trabajo, el tratamiento de esta primera obra acerca de bibliotecas y publicaciones digitales fue hecho tratando de estudiar y explicar la biblioteca digital desde un enfoque reduccionista, como un ente aislado o, cuando mucho, como un grupo de ellas dentro de una organización.

Bajo este primer enfoque, el objeto de estudio fue la biblioteca digital en sí misma, sus partes y sus interrelaciones: las preguntas pertinentes al respecto fueron: ¿qué es exactamente una biblioteca digital y qué rasgos la definen? ¿de dónde viene? ¿existe un sólo tipo de ellas o cuáles son sus variantes? ¿qué partes la conforman? ¿cuáles son sus servicios y colecciones y cómo se desarrollan ambos? ¿cómo es el mundo de las publicaciones digitales que se integra a sus colecciones? ¿cuáles son las semejanzas y las diferencias entre ellas y una biblioteca tradicional? ¿cómo se diseña y construye una biblioteca digital? ¿qué esperan los usuarios de una biblioteca de este tipo?

De esta forma, y teniendo como centro la biblioteca digital, he tratado de analizar minuciosamente todas y cada una de sus partes principales, tratando de explicar sus interrelaciones y de descubrir sus características pensando en un lector profesional de la bibliotecología, para ayudarle a comprender, de la manera más precisa posible, el origen, la evolución, el estado del arte y las tendencias de la biblioteca digital, y el cómo concebir, diseñar y construir una biblioteca digital, disectando cada una de sus partes y sus interrelaciones. He tratado de mostrar cómo se va construyendo una colección digital, con material digital original o digitalizado, y de explicar cuándo, cómo, y por qué debe digitalizarse tal o cual material, y las ventajas y desventajas de

hacerlo; qué implicaciones legales tiene el uso o copia de los materiales digitales y cómo el bibliotecario actual puede desenvolverse adecuadamente en ese entorno. Como hemos mencionado también, este enfoque no ha cuestionado en sí la naturaleza social de la biblioteca, ni la necesidad o pertinencia de su existencia, consolidadas por más de cuatro milenios de historia.

Como resultado de ello podemos ir estableciendo las siguientes conclusiones:

- El concepto de biblioteca digital no es simple; de hecho, varía dependiendo de la comunidad que pretende definirlo. Se trata de un término que además ha venido evolucionando por varias décadas y al que hemos denominado biblioteca electrónica, virtual, sin paredes, ciberteca, de medios, etcétera, hasta desembocar en el concepto más actual, el de biblioteca digital, el cual abarca a los demás le agrega un contexto más actual de acuerdo con el estado del arte de hoy en día.

Podemos agregar además, como lo resume Clifford Lynch,⁸⁴ que “La Internet no es la biblioteca mundial de la era digital. Esta afirmación no se sostiene más allá de una afirmación casual. La Internet, y en particular su colección de recursos documentales conocidos como la World Wide Web, no fueron diseñados para sostener la publicación organizada ni la recuperación de información como lo fueron las bibliotecas. Ha ido evolucionando como un caótico repositorio de la ‘publicación electrónica’ mundial, pero no es una biblioteca digital”. Extiendiendo esta conclusión con el corolario: “Las bibliotecas digitales están contenidas dentro de la *Internet* y por tanto son parte de los recursos documentales de la red mundial. La *red* como un todo no es una Biblioteca Digital”.

Como definición, la mejor que encontré desde mi punto de vista fue la del *Taller de planeación de Santa Fe acerca de medio ambientes de trabajo de conocimiento distribuido* [Santa Fe Workshop on Distributed Knowledge Work Environments, 1997] donde se estableció el consenso, cada vez más aceptado a nivel mundial, de que

84 Lynch, Cliffor d. 1997. “*Searching the Internet...* Op. Cit.

“el concepto de biblioteca digital no es simplemente el equivalente de colecciones digitalizadas con herramientas de manejo de información. Es más bien un ambiente digital para integrar colecciones, servicios y personas en apoyo a un ciclo vital de creación, disseminación, uso y preservación de datos, información y conocimiento. Las oportunidades y retos que motiven ulteriores investigaciones acerca de las bibliotecas digitales deben asociarse a esta amplia visión del ambiente que representan estas bibliotecas”.

Y como la mejor caracterización-definición de biblioteca digital para la comunidad bibliotecaria podemos establecer que sin duda, la mejor de todas es la del *Taller de UCLA-NSF* [UCLA-NSF, 1996]. De acuerdo con este foro se establece que:

“Las bibliotecas digitales son un conjunto de recursos electrónicos y capacidades técnicas asociadas para la creación, búsqueda y uso de la información. En este sentido son una extensión y mejora de los sistemas de almacenamiento y recuperación que manipulen datos en cualquier medio —texto, sonido, imagen estática o dinámica— existentes en redes distribuidas. El contenido de las bibliotecas digitales incluye datos, metadatos descriptivos y metadatos de hipervínculo. Las bibliotecas digitales están construidas —colectadas y organizadas— por una comunidad de usuarios, y las capacidades funcionales de la biblioteca apoyan las necesidades de información y costumbres de esa comunidad. Son componente de esa comunidad en la cual grupos e individuos interactúan unos con otros, utilizando recursos y sistemas de datos, información y conocimiento. En este sentido ellas son una extensión, mejora e integración de una variedad de instituciones de información en lugares físicos donde los recursos son seleccionados, colectados, organizados, preservados y accesados en apoyo a comunidades específicas de usuarios. Estas instituciones de información incluyen, entre otras, a bibliotecas tradicionales, museos, archivos, y escuelas, extendiendo sus servicios a estos lugares, y además a aulas, oficinas, laboratorios, hogares y áreas públicas”.

- Los aspectos más importantes a considerar en el diseño e implementación de una biblioteca digital se establecieron con detalle en el apartado 1.4. En forma concreta éstos son:
 - Enfoque de las colecciones.
 - Enfoque de la tecnología.
 - Enfoque de los usuarios.
 - Enfoque de la organización documental.
 - Enfoque legal.
 - Enfoque social.

Dentro de los elementos analizados en cada uno de estos enfoques, vale la pena resaltar las siguientes conclusiones:

- De acuerdo con el consenso de los principales grupos de especialistas a nivel mundial, después de varios años de observar el desarrollo de las tecnologías para digitalización: sus capacidades, costos, etcétera, el consenso es que "la digitalización es una excelente forma de maximizar el acceso a materiales; es decir, con fines de distribución es una formidable herramienta." Con fines de preservación de documentos a largo plazo el contexto tecnológico todavía no está maduro y debe usarse **con cautela** y siempre y cuando se tenga algún "respaldo" de otro tipo.
- Es sumamente importante que toda biblioteca digital establezca con precisión sus criterios para digitalizar las colecciones de sus documentos que se encuentran en soportes tradicionales, con objeto de evitar errores, inconsistencias y proyectos sin sustento a largo plazo. Esta definición de criterios permitirá clarificar y establecer a la vez aspectos técnicos, como son los formatos que se utilizarán en la digitalización, y dependiendo de sus fines de preservación y/o distribución; tipos de compresión, resolución, etcétera; aspectos de tipo legal; "uso ético", etcétera. Muy importantes son también los "veintidós principios" para el desarrollo ideal de una colección digital. Para más detalle véase el apartado 2.3.

En cuanto a las publicaciones originalmente digitales que son distribuidas por las bibliotecas digitales, se tienen como las principales conclusiones las siguientes —en lo relativo a revistas académicas—:

- Las revistas digitales académicas presentan en efecto una serie de ventajas atractivas para los usuarios; por ello su uso sigue incrementándose cada vez más en las bibliotecas, a nivel mundial. Por lo mismo, cada vez es menor la proporción de bibliotecarios que siguen manteniendo colecciones de revistas exclusivamente de papel.
- Muchos bibliotecarios manejan todavía con cautela este cambio, y mantienen una proporción muy alta de suscripciones en los dos soportes.
- A nivel general, en Norteamérica y Europa, el motor principal del cambio a la versión electrónica proviene de los usuarios y académicos; en América Latina el motor principal son las restricciones presupuestales.
- El modelo de *Open Access* es medianamente popular en E.U.A. y la Unión Europea; no así en bibliotecas de América Latina.

En lo relativo a libros electrónicos:

- El libro electrónico, *e-book*, ha ido desarrollándose también en los últimos años, aunque a un ritmo más lento que el de las revistas académicas. Diversos motivos han ido retrasando su auge al mismo nivel que otro tipo de documentos digitales porque las soluciones y propuestas tienen ventajas sólo en ciertos aspectos. Estos motivos pueden agruparse en tres grandes grupos:
 - Primeramente, los problemas relativos a los dispositivos; carecemos de uno que reúna realmente *todas* las características inherentes a un libro en papel: capacidad, comodidad, costo, portabilidad, confiabilidad, etcétera, que *en conjunto* tiene un libro. Estamos todavía en espera del *dispositivo perfecto*.
 - El segundo grupo de problemas es el que los editores y distribuidores encuentran que se deriva de la seguridad y control de las obras que se venden, así como en algunos de los

mecanismos para control de venta y versiones. Algunas soluciones y medidas, como las *marcas de agua electrónicas* o filigranas, encriptado, temporalidad, etcétera fueron aquí analizadas.

- El tercer grupo de problemas es el derivado de la estandarización. Desde un principio y por bastante tiempo, la edición de *e-libros* ha sido anárquica y de difícil normalización, ya que cada editor ha asumido los formatos, metadatos, estilos, patrones y dispositivos que ha querido. Sólo recientemente se han empezado a introducir y aceptar algunos estándares internacionales como el OEBPS 1.2 —*Open eBook Forum Publications Structure Specification*— versión 1.2 y posteriores, que integran lo mejor de la experiencia adquirida en este tema y prometen empezar a hacer alguna luz en lo relativo a estándares de libros electrónicos.
- Con respecto a otras publicaciones electrónicas encontramos que se van consolidando cada vez más como parte de las colecciones bibliotecarias digitales; entre ellas las más observadas son: Imágenes digitales, audio digital, video digital, bancos de datos de todo tipo y sitios *Web*.

Como conclusión final en lo tocante a publicaciones electrónicas, se observa que ya existen bibliotecas que poseen algunos de los siguientes tipos de colecciones de publicaciones digitales: libros y revistas de todo tipo; atlas, enciclopedias, diccionarios, partituras, mapas, periódicos, semanarios, manuscritos, catálogos, encuadernaciones, iluminaciones, impresos, dibujos y fotografías, archivos, filatelia, patentes, grabaciones digitales de música, naturaleza, narraciones, teatro, literatura, lenguas, dialectos, etcétera.

Con respecto al registro documental, es de particular interés subrayar que tan importante es la creación y desarrollo de colecciones —*ahora digitales*— como el desarrollo de mecanismos que permitan su recuperación y distribución. De nada sirve crear copiosas cantidades de información si ésta no puede ser localizada cuando se requiere o si se deben invertir enormes cantidades de tiempo en decantar información no relevante para nuestra actividad. La información para el estudio,

la investigación, el trabajo o la capacitación debe estar a unos cuantos golpes de tecla cuando se requiera. Debe ser veraz, oportuna, pertinente y suficiente; sus premisas son la organización, la recuperabilidad, la autenticidad y la invariabilidad. De hecho, ésta ha sido y sigue siendo la diferencia esencial entre una verdadera colección documental y una masa amorfa de información. No hay colección digital sin mecanismos de recuperación.

Es un hecho indubitable que el acumulamiento de experiencias por parte de la bibliotecología, lo preciso de las metodologías y el rigor de los procedimientos lograron mucho en el control documental. El formato MARC, las normas ISBD, Las Reglas Angloamericanas de Catalogación AACR 2 nos han llevado mucho más cerca del Control Bibliográfico Universal. Pero la siempre creciente expansión del número de documentos por controlar y la enorme variedad de ellos impone una búsqueda continua de nuevas herramientas cada vez más potentes para contender con la tarea. Las computadoras han sido de gran ayuda para el procesamiento de esta información, pero trajeron a su vez nuevas variedades de documentos, —*los documentos digitales o electrónicos*— y en cantidades tan numerosas que nuevamente incrementaron el problema hasta alturas impredecibles.

Ha habido siempre un problema de balance en el registro documental desde que éste se formalizó: los mejores registros documentales son los efectuados por profesionales de esta actividad siguiendo precisos métodos, auxiliándose con extensas y poderosas herramientas documentales, pero consumen mucho tiempo y recursos humanos calificados, requieren insumos caros, y no hay manera de seguirle el paso a la producción documental. Por otra parte, se han desarrollado mecanismos diseñados para un personal mucho menos entrenado, con sólo unos cuantos elementos prácticos de registro documental, y tales mecanismos son económicos, rápidos y fáciles de hacer y están al alcance de casi cualquier persona o institución; pero la calidad de los registros es baja y crea problemas durante la recuperación. Esta otra aproximación al problema permite procesar mucha más información pero conlleva esa falta de precisión mencionada y deja mucho material sin recuperar, de tal forma el acceso resulta inadecuado. Las preguntas son por tanto: ¿se procesa bien y con calidad sólo una parte del material

documental o se trata de procesar todo con poca calidad? ¿cuál material electrónico vale la pena registrar?

Obviamente ambas posturas han creado debates desde hace muchísimo tiempo, y ha habido defensores de la calidad sobre la cantidad, sabiendo que mucho se quedará fuera y por otro lado hay quien defiende que es mejor que todos los materiales estén procesados, aunque sea pobremente, en lugar de dejarlos sin proceso. La respuesta a este debate no es sencilla y depende del análisis de muchos elementos como los entornos, contextos, recursos y tiempos de quien se hace la pregunta. Han surgido posiciones fundamentalistas en ambos extremos y ha habido quienes buscan algo que optimice ambas soluciones y se ubique en algún punto medio.

Como resultado de ello han surgido las nuevas herramientas para el control documental que son de distinta naturaleza, potencia y alcance, y que han venido a sumarse y no a sustituir a las antiguas herramientas ya mencionadas. Como todos los utensilios, se han creado para distintas tareas y funciones, y por ello el adecuado debe ser seleccionado para la tarea apropiada. Como hemos visto, a muchas de estas nuevas herramientas se les ha llamado *metadatos*, aunque el término, formalmente hablando, abarca perfectamente bien tanto a una ficha catalográfica AACR2 impresa sobre cartón como a una ficha MARC, como ya hemos establecido. Entre estas nuevas herramientas se encuentran las ya conocidas de MARC, Z39.50, etcétera, si bien en nuevas versiones, como el MARC 21 o el MODS —*Metadata Object Description Schema*—. A ellas se han agregado el “Código del Núcleo de Dublín” —*Dublin Core*— y los lenguajes de marcado: SGML, HTML y el más nuevo, poderoso y promotor XML. Las nuevas definiciones de la especificación OEBPS 1.2 —*Open eBook Forum Publications Structure Specification*— para definir formalmente una máquina de libros electrónicos, etcétera.

Bajo mi perspectiva, las conclusiones más importante en lo tocante al control documental son las siguientes:

- El registro documental organizado es fundamental para toda colección, sea ésta digital o no. Nunca deben crearse colecciones digitales sin un mecanismo de registro y recuperación documental.

- Es necesario establecer el nivel del registro documental que sea el más apropiado a la naturaleza del material documental. Colecciones de objetos muy elementales: recetas de cocina, catálogos de venta, colecciones personales, etcétera, tendrán suficiente con mecanismos de registro muy sencillos. Conforme la importancia y el tamaño de la colección sean mayores, el Núcleo de Dublín u otros esquemas parecidos serán más adecuados. Las colecciones formales de libros, revistas, mapas, videos, etcétera, deben utilizar las mejores herramientas de registro: Registros ISBD-MARC-AACR2, OEBPS 1.2, MODS y esquemas XML.
- Las nuevas herramientas de registro mencionadas vienen a sumarse y no a sustituir a las anteriores.
- No existe ningún tipo de metadato que sea la *bala de plata* que nos resuelva todas las situaciones de registro documental para todas las colecciones. Así como para algunas de ellas creábamos fichas AACR2 y para otras índices o resúmenes, o fichas analíticas, hoy es responsabilidad del bibliotecólogo saber escoger una mezcla de herramientas adecuadas que le brinden un balance adecuado entre precisión, velocidad de registro y de recuperación, economía, tamaño de la colección, interfaces, etcétera, para todas y cada una de sus colecciones digitales.
- Hay personas y organizaciones que defendieron ferozmente la ortodoxia del registro documental sólo para ver finalmente cómo eran rebasadas vertiginosamente por la variedad y abundancia documental, lo que resultó poco útil para los usuarios. Hubo quien pensó que ya no era necesaria ninguna regla de registro, o que las reglas podían ser absolutamente laxas, y que a “fuerza bruta” de procesamiento de computadora toda la información podía ser recuperada. Ahí están como prueba algunos de los “buscadores” de Internet, quienes recuperan un verdadero océano de basura documental donde encontrar lo relevante es tarea imposible, lo cual resulta poco útil para los usuarios. La mejor combinación a la fecha sigue siendo un buen buscador combinado con una estructura formal de registros documentales.
- Será necesario seguir insistiendo y profundizando en ese delicado balance entre calidad y precisión contra rapidez y economía.

Habremos de continuar la investigación bibliotecológica e informática buscando todavía mejores soluciones a estos problemas. Algunas cosas apuntan ya como mejoras al respecto, pero ciertamente estamos todavía lejos de la solución definitiva.

Respecto a las publicaciones electrónicas y los derechos de autor y de propiedad podemos establecer lo siguiente: primero, como señalamos, cabe resaltar que el concepto de derecho de autor como concepto jurídico es diferente del *copyright*. El primero proviene del derecho romano-canónico, típico de nuestros países iberoamericanos, y tiene que ver más con el aspecto subjetivo del acto de creación de una obra por un autor dado; el derecho moral consagrado en estas legislaciones tiene su origen en el humanismo y la revolución francesa, de donde se extiende a España y de ahí a los países latinoamericanos y a algunos otros de Europa. Su esencia reside en que la obra forma parte integrante de la personalidad del autor y que es una creación de su espíritu; el fruto de su pensamiento, de manera que no puede ser disociado enteramente de aquél, por lo que incluso cuando el autor ha cedido sus derechos patrimoniales sobre la obra ésta continúa asociada a él y en cierta medida bajo su dependencia. *En resumen, es un derecho del ser*. En contraste, el *copyright* proviene del derecho consuetudinario anglosajón —*common law*—. Ahí, los derechos morales no son primordiales y su esencia está más intrínsecamente ligada al derecho de copiar o explotar una obra; es decir, su enfoque es más comercial, *es un derecho del tener*. En las legislaciones actuales se conocen como derechos *morales y patrimoniales*.

Realmente los bibliotecarios están entonces más relacionados con el derecho de copia, esto es, los derechos patrimoniales que tienen que ver con la explotación y copia de una obra dada. En particular, y dentro del ambiente de las bibliotecas digitales es necesario tener muy claro qué se puede copiar y en que circunstancias. El advenimiento de las publicaciones electrónicas ha ido transformando notoriamente el contexto de los derechos de copia inherentes a las publicaciones y esto no puede simplemente extrapolarse a lo que anteriormente estaba establecido para publicaciones en papel, ya que ello induciría a caer en

algunos errores en el manejo adecuado de los derechos de copia en una biblioteca que contenga o acceda a ese tipo de publicaciones.

Las publicaciones que se pueden copiar sin restricción de derechos patrimoniales o de copia son:

- Las que se encuentran en el dominio público, ya sea por su propia naturaleza, porque sus derechos han prescrito o porque así lo ha decidido expresamente el autor. En este sentido se ha hecho un análisis detallado que el bibliotecario profesional debe estudiar con cuidado, ya que las prescripciones cambian de acuerdo con las leyes internacionales o las de cada país y se mueven en función de extensiones y prórrogas establecidas sobre las obras.
- Las siete excepciones permitidas en la ley mexicana en su capítulo II, artículo 148. Véase apartado 5.1.1.

La falta de precisión en las legislaciones autorales con respecto a nuevas variantes que contemplen de manera adecuada las características de las publicaciones digitales introduce serias lagunas y riesgos enormes para la existencia misma de las bibliotecas y de sus usuarios y deben ser vigiladas y enmendadas por los responsables de esas bibliotecas. Es de capital importancia buscar modelos que restablezcan el equilibrio de la protección a los autores junto con el derecho de los usuarios a la información. Es absolutamente vital rescatar los principios de jurisprudencia de las legislaciones mundiales al respecto de que: “el propósito del ‘copyright’ es promover el progreso del conocimiento dando al autor de una obra un incentivo económico para crear nuevas obras”. La fallida “Acta del Milenio” lo prevé aunque no lo cumple, pero es sumamente importante rescatarlo: “nuestro mandato debía evaluar los efectos de las enmiendas hechas por el título I de la DMCA y el desarrollo del comercio electrónico y las tecnologías asociadas y emergentes sobre la operación de las secciones relativas al título [...] expresamente, este informe se concentra en tres propuestas que nos fueron presentadas durante nuestras consultas: la creación de una *‘doctrina de la primera venta digital’*; la creación de una excepción que *‘permita crear ciertas copias incidentales temporalmente*

y la ampliación de excepciones a el archivado de copias de programas de cómputo”. (El subrayado es mío).

Como ha podido observarse, si bien algunas legislaciones han extendido las protecciones de sus leyes de derechos de autor mucho más allá de lo usual en materiales impresos, también puede observarse que existen ya algunos contrapesos que buscan permitirles a las bibliotecas continuar con su labor, conscientes que de otra forma corren el riesgo de ser asfixiadas por estas disposiciones. “Por siglos hemos desarrollado y arraigado la idea de que si nos gusta algo que compramos para leer, podemos recomendarlo, prestarlo o regalarlo a un amigo o colega, y además si decidimos que ya no deseamos esa obra, podemos tirarla o venderla. Estos principios son tan importantes en nuestro quehacer intelectual que debemos tenerlos presentes siempre y encontrar la manera de trasladarlos al entorno de las publicaciones digitales” [Soules, A. 2002].

Esta preocupación de que los derechos de los usuarios no pueden estar desbalanceados respecto de los derechos de los autores, ha sido recogida claramente por las bibliotecas y sus organizaciones. La Federación Internacional de Asociaciones e Instituciones Bibliotecarias (IFLA) establece claramente estas posturas en su documento *Position Paper on Copyright in the Electronic Environment* sancionado por sus comités en 1996 en la 62ª Conferencia Mundial de esta organización. [International Federation of Library Associations and Institutions, 1996].

El derecho de poseer físicamente un ejemplar de una obra por parte de un comprador ha sido consagrado en todas las legislaciones del mundo durante casi dos siglos. Por ello podemos poseerlo indefinidamente, regalarlo, prestarlo, venderlo, subastarlo, anotarlo, transportarlo, leerlo en voz alta e inclusive destruirlo. Esto se ha vuelto un derecho intrínseco a la naturaleza del que posee un libro u obra semejante. Las bibliotecas y sus colecciones de libros o revistas encajaban perfectamente en este esquema y todo estaba en equilibrio.

Con el advenimiento de los mecanismos de copia masiva al alcance del público y de las bibliotecas este equilibrio empezó a romperse. La fotocopia en los impresos y la grabadora en los registros sonoros permitieron al público y a las bibliotecas obtener mucha mayor cantidad de páginas u obras. A partir de ese momento los editores arremetieron

contra el derecho de copia que había estado intrínsecamente ligado a un lector y que no había sido cuestionado en los siglos anteriores. Las bibliotecas pagaron caro el delito de *lesa copyright* por el servicio que les hacían a sus usuarios; los precios de las suscripciones y otros materiales se vieron incrementados para este tipo de institución, dado su nefando pecado de obtener copias para el público, y de algún modo el equilibrio volvió a ser reestablecido, a pesar de que, como ya hemos mencionado, las bibliotecas sólo aplican la esencia del derecho de copia consagrado en esa original premisa constitucional contenida **en la mayoría** de las legislaciones autorales del mundo: “el propósito del ‘copyright’ es promover la ciencia y las artes útiles a través de la difusión del conocimiento” y que “el autor se beneficia de la distribución ya que ello disemina su obra, creando además de nuevo conocimiento, nuevas ventas”.

Con la llegada de las publicaciones electrónicas y la posibilidad de las bibliotecas y del público de efectuar copias muy fidedignas y masivas de las obras, es necesario poner en contexto nuevamente las legislaciones mundiales para que el equilibrio sea restablecido. Es cierto que el público se vuelve ahora un riesgo potencial para el editor en la medida que puede copiar y distribuir masivamente una obra y por ende debe ser controlado, pero también es verdad que el público no puede ser tratado y restringido partiendo del supuesto de que **siempre** se convertirá en otro editor comercial. El público ha creado derechos y costumbres de cómo usar una obra y a obtener sus *copias incidentales temporales* —como ha sido ratificado nuevamente en la DMCA— a lo largo de siglos, y es imposible ahora simplemente pretender que cuando se paga por una obra sólo se adquiere el derecho de ver esa obra por un periodo finito de tiempo, o en un sólo lugar, y que es exclusivamente para sus ojos y no puede ser compartida en modo alguno. Eso va en contra de la naturaleza hombre-libro. Ningún modelo comercial, tecnológico o legal de alta restricción entre los establecidos hasta la fecha parece prometer algo real a futuro. El equilibrio entre el derecho de comercializar de unos y el derecho de copiar de otros debe ser reestablecido en la era digital; la fórmula debe ser de ganancia para ambos; ninguna otra fórmula funcionará a la larga: *todos los extremos son malos*.

En resumen, pienso que si los editores rescatan los principios y el espíritu básicos que han regido los derechos de copia y los modelos de negocio de algunas industrias editoriales, no tan sólo la del libro, y estos derechos son replanteados con honestidad podrían encontrarse mejores soluciones para el mundo digital; algunos de ellos ya se han mencionado y vienen de origen desde la primera legislación autoral: los autores y editores deben protegerse de otros editores sin escrúpulos que lesionan sus intereses; ellos son los antagonistas y tienen todo el derecho de luchar contra ellos. Sólo cuando un miembro del público decide convertirse en “editor” ilegal se vuelve un antagonista: considerar al público en general siempre como otro editor potencial restringiéndole absolutamente sus prerrogativas de copia “legal” o “ética” va contra la esencia misma que creó el libro y la creación y divulgación del conocimiento. Ningún modelo que vaya en contra del restablecimiento del balance entre el legítimo derecho de copiar y poseer de los usuarios y el legítimo derecho de comercializar y obtener ganancias de los autores y editores funcionará. Cualquier cosa orientada hacia uno de los extremos estará destinada al fracaso, y escindirá cada vez más el espacio entre editores y usuarios, con el beneplácito de quienes detentan esas obras para obtener un lucro indebido. Usuarios y editores deben plantearse mutuamente el paradigma: “si tú me ves como tu socio, seré tu socio; si tú me ves como tu enemigo, seré tu enemigo”.

Ante la falta de legislaciones adecuadas y modernas a este respecto, existe la tendencia mundial, a utilizar las denominadas normas aceptadas para el “uso ético” —*fair use*— de la información, pensando en las aplicaciones de docencia e investigación relacionadas con las bibliotecas. Se destaca el hecho de que con este enfoque se permite la reproducción limitada de obras protegidas con fines de investigación, crítica, reportaje, parodia y docencia —incluyendo varias copias para uso en el aula—. Los cuatro principales parámetros para determinar si un documento dado se copia bajo el principio del *uso ético* y no infringe la ley son:

- El propósito y carácter del uso, que considera si éstos son de naturaleza comercial o con fines educativos no lucrativos.

- La naturaleza del trabajo protegido.
- La cantidad y relevancia de la porción copiada, en relación con la obra en su conjunto.
- El efecto de ese uso sobre el mercado o el valor potencial de la obra copiada. Véase apartado 5.2.2 para más detalle.

No hay absolutamente ninguna receta que permita realizar una copia cien por ciento exenta del riesgo de una reclamación, sobre todo en el ámbito internacional, pero los criterios del *uso ético* tienden a ser más y más utilizados como un estándar *de facto* y su ventaja enorme es su aceptación internacional. Por ello, el número de bibliotecarios que se rigen por él es cada vez mayor en el medio de la información digital.

Otro de los mecanismos que han estado tomando auge en la distribución de material dentro de la red, es el denominado principio del *copyleft*. El propósito del *copyleft*, y de ahí su nombre, era originalmente el de mantener libre a una pieza de *software* dada, y evitar así su privatización. Bajo este principio, la declaración del *copyleft* en una obra es en apariencia la misma que la del *copyright*, e informar el nombre del autor y de su calidad de creador o sobre propiedad intelectual de la obra, sólo que en el *copyleft* el autor autoriza explícitamente a los usuarios a agregar, modificar, mejorar, adaptar, etcétera, los trozos de código que considere pertinentes. Este mecanismo permite además obtener copias de la obra y distribuirlas libremente a discreción, pero lo más importante de todo es que agrega una restricción, **la de agregar restricciones propias** a los subsecuentes usuarios o autores. Esto es de singular importancia, ya que esta restricción de restringir, paradójicamente, se vuelve la clave de una distribución libre. De esta forma, cada usuario es libre de distribuir la obra, pero no puede imponer restricciones adicionales a su distribución. Ese nuevo usuario puede modificarla, señalando inclusive dentro de la obra la parte que es de su autoría, pero no puede restringir la capacidad de otros a modificarla a su vez o distribuirla. De esta manera se garantiza que la obra no podrá ser llevada al régimen propietario más adelante por alguna persona o institución malintencionada. Aunque el concepto del *copyleft* ha sido utilizado como un mecanismo de distribución de *software*, ha

probado ya sus bondades y ganado muchos adeptos; por ello ha empezado a extenderse a otro tipo de obras, como las musicales o a textos, sobre todo aquellos que viajan a través de la red mundial. Cada vez es más frecuente encontrarse con algún tipo de documento u obra de texto o artística que establece su esencia de *copyleft* y mantiene así su capacidad de circular ampliamente y sin restricciones por la red.

Finalmente, como una idea derivada de estos conceptos del *uso ético* y del *copyleft* encontramos otra instancia interesante denominada *Creative Commons*. En esencia, es una metodología que permite al público en general que desea publicar una obra en *Internet* —texto, música, imagen, video, etcétera— seleccionar al instante en su sitio *Web* y sin ayuda de asesoría legal una serie de diversas *licencias tipo* que ya han sido redactadas por expertos en donde el autor selecciona ciertos derechos que él quiere otorgar a quienes lo consulten, y ciertos derechos que él se reserva. Las combinaciones de derechos son muy numerosas y permiten que el autor haga una cesión muy pormenorizada de lo que permite y lo que no. De esta forma, quien encuentra el documento en la *Web* y desea utilizarlo sabe exactamente lo que el autor le permite hacer y lo que no, sin necesidad de contactos o permisos adicionales. Esta “licencia” queda adosada al instante a su documento electrónico de tres formas: legible al público en general, en redacción legal y en formato de máquina legible por computadora por un *navegador*. Esta metodología ha sido adoptada ya en veinte países y está en vías de estudio en otros tantos más porque simplifica enormemente el otorgamiento y retención de derechos en publicaciones electrónicas, sin sustituir las leyes de derechos de autor ni el dominio público; ya probado ser ya un excelente complemento a las disposiciones legales en el mundo de las publicaciones electrónicas.

- Con respecto al perfil del bibliotecario profesional para las bibliotecas digitales, podemos concluir lo siguiente:

Las bibliotecas digitales requieren de bibliotecarios digitales. Las colecciones digitales deben ser planeadas, seleccionadas, adquiridas, organizadas, distribuidas y preservadas. Los servicios digitales deben ser diseñados, construidos, operados y mantenidos. Si bien en términos

generales, los requerimientos para crear y operar una biblioteca digital son básicamente los mismos que para una biblioteca tradicional, las semejanzas terminan allí. Organizar un conjunto de colecciones y servicios digitales representa una serie de acciones y habilidades cotidianas que difieren de las requeridas para hacerlo con sus correspondientes tradicionales. Los bibliotecarios profesionales que dirigen bibliotecas digitales deben tomar decisiones y realizar acciones para las cuales probablemente no fueron entrenados a fondo durante su carrera profesional. Si a esto sumamos el rápido cambio en los aspectos tecnológicos resulta que lo que se aprendió recientemente queda rebasado en poco tiempo. El perfil entonces es “saber ser y saber hacer puesto de manifiesto en actitudes, conductas o comportamientos en situaciones en las cuales pueda encontrarse un profesional”. Existen varios tipos de perfiles: perfil ocupacional, perfil de competencia, perfil de personalidad, perfil teórico o perfil ideal y perfil profesional.

Ese perfil profesional del bibliotecario debe involucrar no sólo el perfil ocupacional sino también expectativas de desempeño acordes con las necesidades del medio y una visión acertada de los cambios tecnológicos y sociales que estamos viviendo. En los estudios relativos a perfiles profesionales, hay varias maneras de dividir los perfiles; una de ellas ha sido dividir en dos grandes vertientes la caracterización del personal, estableciendo por un lado un enfoque hacia los atributos personales, es decir, las características innatas de un profesional, y por el otro los atributos profesionales, o sea las características aprendidas durante su formación. Existen por supuesto variadas subdivisiones o agrupamientos de las características según el enfoque de quienes han realizado estos estudios.

Existen ya definiciones de perfiles deseados para el profesional actual de la bibliotecología en diversas regiones del mundo; entre las más importantes, tenemos los perfiles establecidos por la *Special Libraries Association* de la unión americana, el de la Unión Europea, y en particular el de los países como España o Francia. En el cono sur, en latinoamérica, hay también muy interesantes definiciones al respecto que pueden servirnos como ejemplos a seguir. En el fondo, no difieren mucho unas de otras. Lo importante de todas ellas es comprender que con estas distintas visiones y caracterizaciones de lo que se espera que

sea un bibliotecario profesional en la actualidad, sea suficiente para poder entender la responsabilidad del perfil que los bibliotecarios *digitales* deben presentar ante la sociedad; en particular ante las organizaciones que los contratan y ante los usuarios a los que atienden; los bibliotecarios profesionales mexicanos no son la excepción.

Por supuesto, esto implica una reflexión conjunta entre los profesionales de la bibliotecología actuales en el campo del ejercicio profesional y aquellos que tienen a su cargo la formación curricular de los futuros bibliotecólogos en las escuelas correspondientes. Ello implica que los planes de estudio, balance de cargas curriculares académicas, etcétera, deberán ser replanteados en algún momento y cada vez más con mayor frecuencia a fin de responder a estos perfiles profesionales.

Roberto Garduño⁸⁵ lo resume así:

“También se advierte que el constante desarrollo de la ciencia y la tecnología en las diversas áreas profesionales ha propiciado una acelerada generación del conocimiento, situación que dificulta su estudio permanente en niveles óptimos de aprendizaje. En este sentido se percibe que los fenómenos surgidos con las tecnologías y los requerimientos del mercado laboral han dado lugar a la formación de profesionales de alto nivel. Al mismo tiempo se ha generado el requerimiento de incidir en una mejor preparación de los recursos humanos, lo cual es propiciado por la orientación hacia lo global, a democratizar la educación, a la constante generación de recursos documentales y a incrementar las posibilidades que otorgan los desarrollos tecnológicos... Frente a los nuevos desafíos la bibliotecología requiere diversificar su campo de conocimiento y desarrollar nuevas alternativas de formación tomando en cuenta los avances del conocimiento en el área”.

Muy vigente queda por tanto la reflexión de Jesse Shera⁸⁶ acerca del estado del arte de la profesión bibliotecaria; a pesar de tener ya quince años me parece que se encuentra cada día más vigente dentro del contexto actual:

85 Gar duño, Roberto. 2004. “*La Educación a Distancia...*”. *Op. Cit.*, pp. 7-8.

86 Shera, Jesse H. 1990. “*Lo que el Bibliotecario...*”. *Op. Cit.*, pp. 201-231.

“De todas las profesiones la del bibliotecario es seguramente la más derivativa y sintética, dependiente sobremanera de las disciplinas más formales para la derivación de su propia estructura teórica y cuerpo de práctica [... esta calidad le ha dado a la bibliotecología una posición estratégica única de liderazgo en la integración del conocimiento humano y podría hacer de la bibliotecología una gran fuerza unificadora, no solo en el mundo del saber sino en el de toda la vida humana”.

La segunda aproximación o enfoque de esta obra, como ya se ha mencionado, fue el holístico, y es el tratado en otra obra *Un Modelo de Planeación de Bibliotecas Digitales Para México*. En este caso el objeto de estudio son **todas** las posibles bibliotecas digitales existentes o potencialmente creables en nuestro país. Las preguntas por tanto fueron otras: en primer lugar, es necesario replantear los problemas desde la función misma de las bibliotecas a nivel nacional. Para ello era indispensable abrir mucho más el horizonte y comenzar con un objeto de estudio mucho más amplio. De ahí, y dado que la biblioteca es una institución de servicio para la sociedad es necesario entonces analizar detalladamente ésa sociedad en la que se inserta: sus características, sus necesidades, su problemática; por supuesto, hablamos de la sociedad mexicana actual. Pero esta sociedad no es una isla en el planeta; es necesario entonces analizar antes a la sociedad mundial para contextualizar dentro de ella a la sociedad mexicana. Y se dice que actualmente la sociedad en la que vivimos a nivel mundial es la denominada *sociedad de la información*. Era necesario entonces empezar a estudiar y definir a ésta así llamada sociedad actual: ¿qué rasgos la definen? ¿en qué es nueva y en qué es igual a otras predecesoras? ¿desde cuándo existe? ¿cuál es su contexto? ¿cuáles son sus tendencias, fortalezas y problemática? ¿cuáles son sus riesgos, injusticias y omisiones? ¿por qué nuestra sociedad se autoadjudica este nombre para definirse? ¿es sociedad de la información o del conocimiento?

*“No es la especie más fuerte la que sobrevive;
ni siquiera la más inteligente: es la que mejor
y más rápido se adapta al cambio y evoluciona”.*

Charles Darwin.

CONCLUSIONES

No pretendo en modo alguno haber agotado todos los temas relacionados a la biblioteca digital. Todo aquel que haya leído completo este trabajo sin duda coincidirá conmigo en ello. El tema es de tal forma extenso que sería imposible afirmar siquiera que se ha tratado medianamente bien. Cada tópico que desarrollaba era una puerta que al abrirla me conducía a un corredor lleno de puertas, y cada una de ellas conducía a su vez a otro corredor en forma infinita; ¿o quizá eran otras galerías, tal como Borges imaginaba en su Biblioteca de Babel?: “el universo —que otros llaman la Biblioteca—, se compone de un número indefinido, y tal vez infinito, de galerías hexagonales [...] Desde cualquier hexágono se ven los pisos inferiores y superiores: interminablemente [...] por ahí pasa la escalera espiral, que se abisma y se eleva hacia lo remoto”. Creo que el lector coincidirá conmigo en que la biblioteca digital es sin duda una Biblioteca de Babel, infinita, compleja, escurridiza, cambiante, pero interesante en extremo por lo mismo y con innumerables secretos aún por descubrir y por explicar. Ahora tengo muchas más respuestas que cuando comencé; por consecuencia tengo muchísimas más preguntas.

Como pudo haberse observado también, las bibliotecas digitales tienen dos *personalidades* o *facetas*; por una parte se parecen mucho, muchísimo, a las bibliotecas tradicionales. No obstante, es obvio que a la vez son diferentes entre sí. ¿Qué tanto son iguales y qué tanto son distintas? Para mí la respuesta está en una palabra: *evolución*. Las bibliotecas digitales no son una revolución; son la evolución lógica de sus antecesoras en un mundo que se mueve hacia otro tipo de entorno, fuertemente influido por una cierta tecnología. Las bibliotecas

tradicionales están evolucionando hacia otras digitales, algunas más rápido que otras, algunas con más éxito que otras; algunas se encuentran en la escala evolutiva mucho más alto y otras se encuentran más abajo; algunas más están en el mismísimo comienzo.

Salvo el no muy frecuente caso de creación de una biblioteca totalmente nueva, sin ningún antecedente y concebida como digital desde su inicio, los encargados de bibliotecas y de sistemas bibliotecarios se enfrentarán por lo general con la responsabilidad de dirigir esos procesos evolutivos en bibliotecas ya existentes. La tarea de los bibliotecarios que se encuentran al frente de estas instituciones es por tanto tomar una biblioteca que se encuentra en un mayor o menor grado de evolución y llevarla a un punto superior de ese camino evolutivo hacia una biblioteca digital. No es un proceso que se logre al cabo de un plazo corto. Por lo general, se trata de procesos que toman varios años, dependiendo de cuánto material se haya acumulado previamente, del tipo de colecciones que se tengan y de que éstas sean susceptibles de convertirse; de los recursos humanos, económicos y tecnológicos que están disponibles para ese proceso, etcétera.

Por tanto, si el proceso es de *evolución* y no de *revolución*, hay un error que cometen frecuentemente muchas instituciones al crear proyectos nuevos de bibliotecas digitales paralelamente a sus bibliotecas tradicionales, —y son proyectos *paralelos* en el más estricto sentido de la palabra: jamás coinciden ni se cruzan— con lo cual desconocen todo el esfuerzo, los recursos, el talento y la experiencia acumulados. Los mejores proyectos de bibliotecas digitales han sido aquellos que capitalizan esos recursos y talentos y los proyectan hacia un nuevo capítulo, digital, de sus bibliotecas. No existe hoy en día la *Digital British Library* que sustituye a la *British Library*, ni la *Digital Library of Congress* que sustituye a la antigua, ni la *Digital Harvard Library* que toma el lugar de la tradicional en su institución, por citar algunas. Todas ellas, enormes, poderosas y reconocidas bibliotecas digitales son la misma biblioteca de siempre, no una paralela; son la misma organización proyectada desde hace varios años hacia un serio camino evolutivo y en el cual se van consolidando como líderes en ese nuevo contexto digital. Nadie puede negar su avance y su evolución, y sin embargo, no son cien por ciento digitales; su camino evolutivo no ha terminado.

Al contrario, se encuentran todavía muy lejos de ese punto final. Los bibliotecarios que las dirigen van caminando en esa dirección evolutiva, aprovechando lo que les heredaron sus antecesores y otros construyendo lo que les dejarán a sus sucesores, ayudando entre todos a esa bibliotecas a andar ese camino evolutivo.

Pero los bibliotecarios que aún no han comenzado no deben esperar expresamente a que los encarguen de dirigir una biblioteca digital. Pueden y deben comenzar ya ese proceso de evolución, a la velocidad que sea posible, sin postergarlo más, dentro de la biblioteca que dirigen hoy en día. Por supuesto habrá que elaborar un plan, y obviamente, se requieren nuevos conocimientos, nuevas habilidades y nuevas actitudes. Es necesario entender lo que es la biblioteca digital para saber hacia dónde nos dirigimos y qué queremos construir. Esta obra pretendió precisamente tratar de apoyar un poco a los bibliotecarios en cuanto a esa planeación, conocimientos y aptitudes, y ayudarlos a entender qué son las bibliotecas digitales.

Umberto Eco explicó alguna vez el significado de su genial y misteriosa última frase de *El Nombre de la Rosa*⁸⁷; básicamente, él quiso decir: «las cosas existen cuando tienen nombre; la sustancia de las cosas proviene primordialmente del nombre, y aunque desaparezcan, queda su esencia en el nombre». Carl Gustav Jung lo dijo también de otra forma: «los nombres invocan; si se pronuncian ciertos nombres las cosas aparecen». He tratado de lograr que el nombre de *bibliotecas digitales* con todo su concepto, con toda la carga y la fuerza, con todo el presente y el futuro que llevan asociados se juntaran en una verdadera esencia y aparecieran. Ojalá y las bibliotecas digitales tengan ahora un poco más de nombre y de significado. Al menos, para mí sí lo tienen.

87 «*Stat rosa pristina nomine, nomina nuda tenemus*» (*La original rosa existió por su nombre, sólo nos quedan los nombres*). Eco, Humberto. 1980. *Op. Cit.*, p. 607.

“Cetera, desunt...”⁸⁸

⁸⁸ “Lo demás, falta...”

GLOSARIO

AACR2. —Anglo American Cataloguing Rules— Reglas Angloamericanas de Catalogación, versión 2.

ACRL Association of College and Research Libraries.

ALA. American Library Association.

Alfabetización (digital, electrónica, informática o tecnológica). Término utilizado para designar el concepto de capacitación básica o mínima del individuo en tecnologías de la información y comunicación.

Alfabetización funcional. Término utilizado para designar el dominio que tiene el individuo en cuanto a la capacidad de obtener beneficios reales para escalar a un mejor nivel de vida gracias al acceso a las tecnologías de información y comunicación.

Algoritmo. Término proveniente del nombre de un célebre matemático árabe del siglo IX, llamado *Al-Juarizmi*, quien escribió un compendio para el cálculo llamado *Kitab al-jabr w-al-muqabala*, en el cual presenta la solución de varias ecuaciones lineales y cuadráticas a través de un nuevo proceso, que consiste en un método para resolver un problema numérico a través de reglas claras, en forma secuencial e iterativa. Posteriormente, a cualquier proceso que siguiera esa aproximación se le denominó en latín con el nombre de su creador, llamándolo “algoritmo”, y de ahí al español.

Ancho de banda. La cantidad de información que puede ser enviada o recibida por unidad de tiempo por un dispositivo de telecomunicaciones dentro de una red, tales como módems, líneas telefónicas, etcétera. Se mide en “bytes “ o caracteres por segundo y sus múltiplos: kilobytes por segundo (kbps), megabytes por segundo (mbps), etcétera.

ARL. Association of Research Libraries.

ASCII. *American Standard Code for Information Interchange.* —Se pronuncia “aski”—. Código creado a fines de los años cincuenta y perfeccionado durante más de una década; a fines de los sesenta tomó la forma que conocemos hoy. En un principio, las computadoras sólo manejaban internamente números. Por su construcción, de naturaleza eminentemente eléctrica, se utilizaron números binarios, ya que es la

representación más simple de un sistema numérico al usar sólo dos componentes “0” y “1”. Ideal para una máquina que tiene cables, energía y todo puede representarlo con “existe electricidad” o “no existe electricidad”. Al principio, los datos eran sólo números, y las instrucciones para manejarlos se daban también con instrucciones en forma de números. Conforme las computadoras evolucionaron se hizo necesario que también pudiesen representar internamente letras, para así ir formando textos más comprensibles para los seres humanos, y después de las letras, los otros caracteres inherentes a un texto —*comas, puntos, signos, etcétera*—. Para representarlos internamente sin romper ese esquema básico binario la solución fue representarlos como una convención de números binarios, arbitraria y previamente establecida, algo así como el código morse. En un principio cada empresa de computadoras estableció su convención y pronto se hizo necesario estandarizarla, por lo que surgió una norma en ese sentido. Ése es el principio del código “ASCII”, el cual nació como una norma norteamericana para homogeneizar las posibles representaciones numéricas binarias de las letras y signos de puntuación. En un principio el código funcionaba sólo con el alfabeto inglés. Luego le fueron agregados otros caracteres con signos diacríticos usados en otros alfabetos occidentales —*letras con acentos, tildes, diéresis, etcétera*—, y de ese modo se convirtió en un código internacional empleado ya no sólo en la unión americana, y que derivó en versiones ISO usadas en todos los alfabetos del mundo occidental. Posteriormente se crearon otras versiones similares para alfabetos griego, cirílico, árabe, japonés, etcétera. Gracias a “ASCII” es posible que lo que tecleamos en cualquier computadora que tenga un teclado “occidental” sea legible en cualquier otra computadora, sin importar la marca de la misma o el sistema que se utilice para enviarlo de forma remota. (Véase también: UNICODE). *ISO 10646. Information technology. Universal Multiple-Octet Coded Character Set (UCS). Part I. Architecture and Basic Multilingual plane. Geneva: ISO [1993-97].*

Blog. Palabra que proviene de la contracción de un término de la red: *Web Log —contenidos personales en la Red—*. En

- esencia, un “blog” es básicamente una revista o serie de artículos que se distribuyen por la red mundial *www*, que tiene la característica distintiva de que su información se actualizan constantemente con nuevos comentarios, datos y vínculos, algunos de ellos se actualizan diariamente; es decir, la información no permanece estática como en una revista “típica”.
- BNF.** Backus Naur Form. Notación para metalenguajes muy utilizada para definir lenguajes de programación desde 1959, y que toma su nombre de sus creadores John Backus y Peter Naur.
- BNF.** Bibliothèque Nationale de France.
- Byte.** Conjunto formado generalmente por ocho dígitos binarios o *bits*, que bajo una convención preestablecida representan un carácter de texto en código ASCII. También se le llama “octeto” o “carácter” y sirve como unidad de medida de almacenamiento en discos, CD, DVD, etcétera.
- Ciber espacio.** Término creado en 1984 por William Gibson en su novela de ciencia ficción *Neuromante* —Neuromancer—. Se define como el espacio de comunicación abierto por la interconexión mundial de las computadoras y por donde fluyen los datos, archivos y mensajes informáticos permitiendo su intercambio por acceso remoto.
- DLTC.** (Canadian) Digital Libraries Technical Committee
- DTH.** Direct-To-Home TV. Sistema de TV enviado directamente al usuario vía satélite gracias a una pequeña antena receptora que hay en el lugar del usuario.
- Dublin Core.** Es un conjunto de quince etiquetas preestablecidas para agregar metadatos a documentos en la Internet con objeto de lograr, de manera muy simple la descripción de recursos documentales destinados a ese entorno de red. El Núcleo de Dublín es el resultado de un consenso internacional multidisciplinario logrado a través de los esfuerzos de una iniciativa que se dio para este efecto. Ha funcionado como un cimiento para una estandarización de registro y descripción bibliográfica de un buen número de documentos disponibles vía la *Internet*. ISO 15836:2003.
- ECIAA.** European Council of Information Associations.
- En-línea.** —*On-line*—. Concepto que establece que existe una conexión “en vivo”; es decir, existente en ese momento entre

dos computadores, o entre una terminal y una computadora, ya sea con una red local o con la Internet.

ENIAC I. *Electronic Numerical Integrator And Calculator*. (computadora e integrador numérico electrónico). Considerada la primera computadora electromecánica en la historia. Comenzó a funcionar en 1946, en la Universidad de Pennsylvania, EUA. Ocupaba 150 m., tenía 18,000 bulbos y pesaba 30 toneladas. Aunque tardaba casi medio segundo en multiplicar dos números de diez cifras, era mucho más rápida que cualquier otra máquina inventada hasta entonces y demostró que los procesos numéricos electrónicos eran posibles, lo que abrió el camino para las computadoras actuales.

Esteganografía. Es el arte o ciencia de comunicar de manera oculta un mensaje, escondiendo o disfrazando la información entre otro conjunto de datos visible, de forma que pase inadvertido ante cualquier inspección sin levantar sospechas. Los primeros documentos que describen el uso de estas técnicas datan de los tiempos de Herodoto en la Grecia antigua. El mensaje microfilmado hasta parecer el punto de una letra en una carta es el arquetipo más extensamente conocido. Hoy en día suele utilizarse para esconder información en todo tipo de archivos tales como fotos, videos o audio. Se puede ocultar un mensaje dentro de contenidos multimedios, mezclando los bits del mensaje original entre los bits del archivo gráfico o de sonido. El archivo resultante será una imagen o archivo de audio totalmente funcional que a primera vista no levanta ninguna sospecha, y del que con el *software* adecuado es posible extraer la información oculta.

FTP —*File Transfer Protocol*—. Protocolo de Transferencia de Archivos vía *Internet*. Como su nombre indica, fue diseñado para transferir todo tipo de archivos entre computadoras. Existen dos tipos de transferencias: la “descarga” —*download*—, consistente en traer y alojar (bajar) un archivo a nuestra computadora desde un servidor remoto; y la “carga” —*upload*— consiste en llevar y alojar (subir) un archivo desde nuestra computadora a un servidor remoto. Existen además dos tipos de FTP muy utilizados: el *FTP anónimo*, en donde se permite a cualquiera que acceda al sitio descargar

un archivo; y el *FTP privado*; en donde es necesario conectarse al servidor utilizando un nombre y una contraseña.

G7, G8. Grupo de los 7. Denominación original de un grupo formado por los países altamente industrializados para realizar esfuerzos cooperativos. En la actualidad son 8 países más la Unión Europea que funciona como un noveno elemento. (EUA, Inglaterra, Francia, Alemania, Japón, Italia, Canadá, Rusia). Se conoce también como G-8 o Grupo de los Ocho.

Gopher . Protocolo desarrollado por la Universidad de Minnesota previo al advenimiento de la *World Wide Web* para organizar y desplegar archivos en servidores en *Internet*. El servidor *Gopher* presentaba sus archivos como una lista jerárquicamente estructurada dentro del mismo protocolo con objeto de encontrar y transferir esos archivos a través de la red. Fue sustituido por los “navegadores” para *Web* .

Handheld. Véase PDA.

HDTV. High Definition TeleVision. Televisión de Alta Definición.

Es un nuevo tipo de TV que provee de mucha mayor resolución de imagen —por lo general el doble— que la TV tradicional, así como pantallas más anchas. No existe un estándar único, varía con cada compañía que lo pretende utilizar.

ICANN —*Internet Corporation for Assigned Names and Numbers*—. La corporación internacional que se encarga de llevar el control mundial de todos los números y nombres de dominio de los sitios en Internet.

Infraestructura de Información. Plataforma de computadoras y telecomunicaciones y los servicios genéricos sobre ellas que soportan las aplicaciones o programas puestos a disposición de los usuarios vía la red.

ISDN. Una red intergubernamental establecida bajo los auspicios de UNESCO con sede en París cuyo objetivo es desarrollar y mantener un registro internacional de publicaciones periódicas, el cual contiene la información necesaria para hacer la identificación y el control bibliográfico, inclusive el Número Internacional Estándar de la Revista (ISSN y el título clave).

JSTOR. Journal Storage Project. Organización no lucrativa creada en 1995 cuyo fin es aprovechar las tecnologías de información para convertir las colecciones de revistas académicas que están en soportes de papel hacia digitales.

KWIC. Tipo de índice permutado en el que el contenido temático de una obra se representa mediante, palabras clave de su título o de otra fuente de información del documento, extraídas de listas de palabras significativas y no significativas. Las palabras clave aparecen en orden alfabético, jerarquizadas de alguna manera en una posición fija, y van precedidas y seguidas por el resto de las palabras del título o fuente.

KWOC. Tipo de índice permutado que varía en su presentación respecto al índice KWIC en el que las palabras clave aparecen como un encabezamiento en línea separada. Bajo cada encabezamiento aparece la totalidad de los títulos, completos o truncados, que contienen la palabra clave de la que se trata.

Metadato. Dato acerca de los datos. Elemento que describe el contenido, condiciones, características, etcétera, de un documento con el fin de definir, identificar, organizar, indizar, filtrar, colocar, preservar, recuperar y administrar ese documento como una parte de un conjunto ordenado de recursos de información electrónicos. Los metadatos pueden crearse para describir los atributos inherentes de un recurso documental cualquiera que éste sea: objeto bibliográfico, —*libro, revista, tesis, etcétera*— registros e inventarios archivísticos, objetos geoespaciales, recursos visuales, —*realia*— y de museos y galerías; programas de computadora, etcétera. Los metadatos pueden presentar diferentes niveles de especificidad, profundidad o estructura. Como ejemplo, la catalogación produce una ficha con metadatos acerca de un libro.

Módem. Contracción de *modulador-demodulador*. Aparato que convierte —modula— las señales de comunicación de una computadora al tipo de señal de una línea telefónica para que puedan viajar a través de ella. Del otro lado, un aparato semejante reconvierte —demodula— la señal nuevamente a su forma original para la computadora destino.

MODS. —Metadata Object Description Schema—. Esquema de metadatos para descripción de objetos. Consiste en un “esquema” XML ya preestablecido por la Biblioteca del Congreso de los EUA para un tipo de documento dado.

NII National Information Infrastructure. Iniciativa creada en los EUA en 1993 cuyo objetivo es proveer la integración de

hardware, software y las habilidades que faciliten el acceso y conectividad masivos entre las personas de ese país, respecto de las ofertas de servicios y recursos de información. Véase también: <http://www.kmbook.com/nii/>

NCOIT&R *National Coordination Office for Information Technology Research & Development*. Agencia coordinadora en los EUA del programa de Infraestructura de Información, el HPCC.

NSF. National Science Foundation. Agencia independiente del gobierno de los EUA establecida en 1950, con el objetivo de promover el progreso de la ciencia en ese país. Muchos proyectos de biblioteca digital en la unión americana son financiados por esta agencia.

Objeto digital. Cualquier entidad documental: texto, imagen, sonido, etcétera, que ha sido codificada numéricamente bajo algún formato de representación digital y ensamblada junto con algún conjunto de metadatos de tal forma que puede ser almacenado, buscado, encontrado, y usado a partir de una colección en un sistema computacional.

OCDE. *Organisation de Coopération et Développement Économiques*. Organización mundial para favorecer intercambios y actividades para la cooperación y el desarrollo económicos.

OEBPS 1.0 — *Open eBook Forum Publications Structure Specification*— Primera especificación consensada internacionalmente acerca de la estructura de publicación de un libro electrónico. En agosto de 2002 este organismo publicó la versión 1.2 de estas recomendaciones, la cual contiene mejoras.

Palm. Véase PDA.

Papel electrónico. Véase: “Tinta electrónica”.

PDA. *Portable Digital Assistant* o Asistente Digital Portátil. Es una pequeña computadora de mano que se opera con un “lápiz” electrónico sobre una pantalla sensible. La mayoría de estos equipos realizan funciones sencillas, como control de agenda, directorio, calculadora, *block* de notas, memorandos, juegos, etcétera. Dependiendo de su capacidad y precio pueden contar con versiones simples de sistemas operativos como el “*Windows CE*” de Microsoft o *PALM OS*, los cuales les permiten extender sus capacidades a proceso de texto, hoja de cálculo, acceso inalámbrico a Internet, etcétera. Permiten

también la inclusión de algún programa que pueda copiar partes de *e-books* a la máquina y leerlos a través de ella.

PyMEs. Sector económico de un país que comprende las “Pequeñas y Medianas Empresas”.

TCP/IP. Es un protocolo que permite la transmisión fiable de *paquetes* de datos sobre redes. Un protocolo es un conjunto de reglas preestablecidas entre dos dispositivos para permitir su enlace. El nombre TCP / IP Proviene de dos protocolos originales de UNIX, el *Transmission Control Protocol* (TCP) y el *Internet Protocol* (IP), si bien en la actualidad existen más de 100 protocolos diferentes definidos en este conjunto. El TCP / IP es la base de la Internet que sirve para enlazar computadoras pequeñas y grandes, provenientes de diversos fabricantes y con diferentes sistemas operativos,.

TIC. Tecnologías de Información y Comunicación. La suma de todo el desarrollo tecnológico humano en computación, informática, redes, telecomunicaciones, etcétera, utilizado para crear, almacenar, intercambiar y usar información en sus múltiples formas (*texto, datos, voz, imágenes, video*). Abarcan equipos de cómputo, programas, servicios en *Internet*, teléfonos fijos y celulares, equipos y servicios de redes, etcétera. Por lo general se asocian a un contexto: TIC para la educación, la salud, el desarrollo, en bibliotecas, etcétera.

Tinta electrónica. Es un material que es insertado dentro de una película para ser usada integrándola a dispositivos de despliegue electrónico; la tinta electrónica es un revolucionario y complejo concepto que implica a la química, física y electrónica y que fue desarrollado para crear este nuevo material. El componente principal de la tinta electrónica consiste en millones de microcápsulas diminutas, aproximadamente del diámetro de un cabello humano. Básicamente, cada microcápsula contiene a su vez múltiples partículas blancas positivamente cargadas y partículas negras negativamente cargadas suspendidas en un fluido claro. Cuando un campo eléctrico positivo es aplicado en la parte inferior de la cápsula, se produce un movimiento de partículas blancas hacia la parte superior de la microcápsula donde se hacen visibles al usuario; esto hace a la superficie parecer blanca en aquel punto. Al mismo tiempo, el campo eléctrico envía las partículas

negras al fondo de las microcápsulas donde aquellas son escondidas. Invirtiendo este proceso, las partículas negras aparecen en lo alto de la cápsula haciendo que la superficie parezca oscura en aquel punto y viceversa. Para formar un despliegue de texto o imagen con la tinta electrónica, las microcápsulas son emparedadas dentro de una película plástica transparente tan delgada como una hoja y que funciona como una capa de circuitería. Bajo una corriente eléctrica predeterminada, esta capa forma un patrón de píxeles que pueden ser controlados por ese dispositivo de despliegue. Esta película puede ser sobrepuesta a casi cualquier superficie como metal, plástico, vidrio, tela, incluso papel, convirtiendo esa superficie en un dispositivo de despliegue.

UNICODE Es el siguiente paso después de ASCII en el desarrollo de códigos binarios de caracteres para escritura o texto. Oficialmente llamado el *Estándar Mundial de Caracteres Unicode*, consiste en un código establecido internacionalmente para el intercambio, procesamiento y despliegue de los textos escritos en la inmensa mayoría de las diversas lenguas del mundo moderno. También apoya el tratamiento de muchos textos clásicos e históricos en varias lenguas. *The Unicode Standard Version 4.0. (2003). The Unicode Consortium. Reading, MA: Addison Wesley. 1504 p. ISBN 0321185781. ISO/IEC 10646.*

Usabilidad: La capacidad de un programa para ser comprendido, aprendido, usado y resultar atractivo para el usuario, en condiciones específicas de uso. ISO - Organización Internacional para la Estándarización.

USB. Universal Serial Bus. Puerto Serie Universal. Puerto de comunicación de computadoras personales con dispositivos externos. Sustituye al RS232C.

REFERENCIAS BIBLIOGRÁFICAS

(Todas las referencias electrónicas han sido verificadas como existentes y exactas hasta el 3 de noviembre del 2005). Están Incluidas las Bibliografías Conjuntas de *Bibliotecas y Publicaciones Digitales y de Un Modelo de Planeación de Bibliotecas Digitales para México* para una referencia total.

American Library Association (ALA). 1983. *The ALA Glossary of Library and Information Science*. American Library Association, Chicago, IL.

An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Author's or Purchasers of Such Copies. Queen Anne of Great Britain, c. 19 (1709). Disponible noviembre 3, 2005 en: http://press-pubs.uchicago.edu/founders/documents/a1_8_8s2.html

Archivo de la Universidad de Salamanca (AUS). Libros de Claus-tros, Libro 239, marzo 11, 1776.

Arias Terry, Ana. 1999. "Demystifying the e-Book - What is it, Where Will it Lead us, and Who's in the Game?" En: *Against the Grain*. Noviembre 1999. Disponible noviembre 3, 2005 en: http://www.against-the-grain.com/ATG_AnaEbook.html

Association of College and Research Libraries. (ACRL). 1998. *A Progress Report on Information Literacy*. March 1998. Disponible noviembre 3, 2005 en: <http://www.ala.org/ala/acrl/acrlpubs/whitepapers/progressreport.htm>

Association of Research Libraries. (ARL). 1995. "Realizing Digital Libraries". En: *Proceedings of the 126th Annual Meeting*. Boston, Massachusetts, May 17-19, 1995. Appendix II. Disponible noviembre 3, 2005 en: <http://www.arl.org/arl/proceedings/126/2-defn.html>

- Association of Research Libraries. (ARL). 2000. *ARL Directory of Scholarly Electronic Journals*. Disponible noviembre 3, 2005 en: <http://www.arl.org/scomm/edir/>
- Ausubel, David, Novak, Joseph. & Hanesian, Helen. 1983. *Psicología Educativa. Un punto de Vista Cognoscitivo*. México:Trillas. 623 p.
- Baca, Murtha (Ed.). 2000. *Introduction to Metadata: Patways to Information*. Getty Research Institute. Página Web del Instituto. Disponible 3 de enero, 2005 en: http://www.getty.edu/research/conducting_research/standards/intrometadata/
- Backus-Naur Form. 2000. En: *FOLDOC: Free On-Line Dictionary of Computing*. Disponible noviembre 3, 2005 en: <http://foldoc.doc.ic.ac.uk/foldoc/foldoc.cgi?query=backus-naur&action=Search>
- Banner, William A. Profesor de filosofía de la Universidad Howards. 1983. *Discurse for the Dedication of the Undergraduate Library*. Septiembre 23, 1983. Disponible noviembre 3, 2005 en: <http://www.howard.edu/library/Special/Excellence@Howard/bannerWA/UGL1983.htm>
- Baseline Focus Group. 1999. *Advantages of Electronic Journals*. Sumario y Reporte Total. Disponibles noviembre 3, 2005 en: <http://www.mimas.ac.uk/sj/baserept.htm>
- Bates, Anthony W. 1995. *Technology, Open Learning and Distance Education*. New York: Routledge.
- Bawden, David y Rowlands, Ian. 1999. Digital Libraries: Assumptions and concepts. En: *Libri*. vol. 49:4, pp. 181-202.
- Beniger, James R. 1986. *The control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, MA: Harvard University Press. ISBN: 0-674-16986-7.

- Bevis King's Video Formats. Disponible noviembre 3, 2005 en:
<http://www.ee.surrey.ac.uk/Contrib/WorldTV/video.html>
- Bibliotheca Universalis. 2004. Página *Web* de la Biblioteca. Disponible noviembre 3, 2005 en: <http://www.culture.gouv.fr/g7/>
- Bibliothèque Nationale de France (BNF). 2004. Página *Web* de la Biblioteca. Disponible noviembre 3, 2005 en: <http://www.bnf.fr>
- Birmingham, William P. 1995. "An agent-based architecture for digital libraries". En: *D-Lib Magazine*. (Junio 1995.)
- Block, David. 2004. "Globalization, Transnational Communication and the Internet". En: *International Journal on Multicultural Societies (IJMS)*, vol. 6:1, pp. 13-28. ISSN: 1564-4901.
- Boomgaarden, Wesley (ed.). 1993, *Staff Training and User Awareness in Preservation Management*. Preservation Planning Program Guide 2. Washington, D.C.: Association of Research Libraries.
- Borgman, Christine L. 1999. "What are Digital Libraries?: Competing Visions". En: *Information Processing and Management*, vol. 35, pp. 227-243.
- Boyce, Peter y Dalterio, Heather. 1996. "Electronic publishing of Scientific Journals". En: *Physics Today* (Enero). American Institute of Physics, p. 42-47.
- Braslavsky, Cecilia. 1994. "Una función para la escuela (...)". En: *¿Para Qué Sirve la Escuela?*. Filmus, Daniel (comp.). Buenos Aires: Norma.
- Braun, Linda W. 2002. "A Librarian by Any Name". En: *Library Journal*. 127(2):46. Disponible noviembre 3, 2005 en:

<http://www.libraryjournal.com/index.asp?layout=articleArchive&articleId=CA191647&display=searchResults&stt=001>

Brazil. Ministry of Science and Technology. 2000. *Information Society in Brazil: Green Book*. Disponible noviembre 3, 2005 en: <http://socinfo.org.br>

British Library. 2004. Página Web de la Biblioteca. Diponible noviembre 3, 2005 en: <http://www.bl.uk/about/policies/digital.html>

Brown, John y Duguid, Paul. 2000. *The Social Life of Information*. Harvard Bussiness School Press. ISBN 0-87584-762-5. 336 p.

Bush, Vannevar. 1945. "As we may think". En: *The Atlantic Monthly*, 176:1 (Julio 1945), pp. 101-108. Disponible noviembre 3, 2005 en: <http://www.theatlantic.com/doc/194507/bush>

Cano, Virginia. 1998. *De Bibliotecario a Gestor de Información : ¿Cambio de Nombre o Nuevas Competencias?* III encuentro de Directores y II de Docentes de Escuelas de Bibliotecología y Ciencias de la Información del Mercosur, Octubre 29 al 31, 1998. Disponible noviembre 3, 2005 en: <http://www.utem.cl/deptogestinfo/extension.htm>

Caplan, Priscilla. 1995. "You Call it Corn, We Call it Syntax-independent Metadata for Document-like Objects". En: *The Public Access Computer Systems Review*, v. 4, no. 6.

Castells, Manuel. 1999. *La Era de la Información: Economía, Sociedad y Cultura*. México: Siglo XXI, p. 375.

Castells, Manuel. 1999. *Internet y la Sociedad Red: Lección Inaugural del Programa de Doctorado Sobre la Sociedad de la Información y el Conocimiento*. Universidad

- Abierta de Cataluña. Disponible noviembre 3, 2005 en: <http://www.uoc.edu/web/esp/articles/castells/print.html>
- Chartier, Roger. 1999. *Cultura Escrita, Literatura e Historia*. México: Fondo de Cultura Económica. ISBN: 968-165974-0. pp. 48, 50.
- Chartier, Roger. 1994. *El Orden de los Libros*. Ed. Gedisa (España). ISBN: 84-7432-498-X. pp. 3-4.
- Cleveland, Gary. 1998. *Digital Libraries: Definitions, Issues and Challenges*. (UDT Occasional Paper #8). IFLANET: Activities & Services. Página Web Oficial de la IFLA. Disponible noviembre 3, 2005 en: <http://www.ifla.org/VI/5/op/udtop8/udtop8.htm>
- Commission on Preservation and Access. (EUA). 1994. *The Digital Preservation Consortium Mission and Goals*. Disponible noviembre 3, 2005 en: <http://www.clir.org/pubs/reports/dpcmiss/dpcmiss.html>
- Collier, Mel. 1997. "Towards a General Theory of the Digital Library". En: *Proceedings of the International Symposium on Research, Development and Practice in Digital Libraries: ISDL'97*. November 18 - 21, 1997. Disponible noviembre 3, 2005 en: <http://www.dl.ulis.ac.jp/ISDL97/proceedings/collier.html>
- Columbia University. 2001. *Selection Criteria for Digital Imaging Projects*. Disponible noviembre 3, 2005 en: <http://www.columbia.edu/cu/libraries/digital/criteria.html>
- Communications of the ACM (CACM). 1995. Association for Computing Machinery. Special Number for Digital Libraries. Vol 38:4. April 1995.

ComputerUser Dictionary. Entrada por *electronic publishing*. Disponible noviembre 3, 2005 en: <http://computer-user.com/resources/dictionary/noframes/nf.definition.html?bG9va3VwPTUyNA==>

Cox, Ingemar J. and Miller, Matt L. 2001. "Electronic Watermarking: The First 50 Years." En: *Proceedings of the IEEE 2001 Int. Workshop on MultiMedia Signal Processing, 2001*. Disponible noviembre 3, 2005 en: <http://www.ee.ucl.ac.uk/~icox/papers/2001/mmisp01.pdf>

Crawford, Timothy. 1992. *Blessing and Curse in Syro-Palestinian Inscriptions of the Iron Age*. American University Studies: Series VII, Theology and Religion, Vol. 120. New York: Peter Lang Publishing. ISBN: 0820-41662-2. 259 p.

Cruz P., Saúl. 2003. "Composición del Presupuesto Asignado a TIC en el Gobierno Federal". En: *Política Digital*. Disponible noviembre 3, 2005 en: <http://www.politicadigital.com.mx/index.php?id=591>

Cyan Worlds. 2001. Página *Web* de la empresa. Disponible noviembre 3, 2005 en: <http://www.cyan.com>

Crawford, Walt. 2000. "Nine Models, One Name: Untangling the e-book Muddle". En: *American Libraries*. 31:8 (September 2000). pp. 56-59.

DARE. Digital Academic Repositories Een Surf-Programma. 2005. Sitio Web oficial del Programa DareNet. Holanda. Disponible noviembre 3, 2005 en: <http://www.darenet.nl/page/language.view/keur.page>

Darnton, Robert. 1993. "Historia de la Lectura", En: *Formas de Hacer Historia*. (Ed. Peter Burke), Madrid. p. 189.

De la Parra, Eduardo. 2004. "Comentarios a las Reformas a la Ley Federal del Derecho de Autor". En: *Revista de Derecho Privado*. Instituto de Investigaciones Jurídicas, Uni-

- versidad Nacional Autónoma de México (UNAM). Nueva época, año III, núm. 8, mayo-agosto 2004. pp. 95-110.
- Deleuze, Gilles y Guattari, Félix. 1987. *A Thousand Plateaus*. Minneapolis: University of Minnesota Press. p. 25.
- Derrida, Jacques. 1976. *Of Grammatology*. Trans. Gayatri Chakravorty Spivak. Baltimore: John Hopkins University Press. p. 8.
- Digital Libraries Federation. 2001. Disponible noviembre 3, 2005 en: <http://www.diglib.org/dlhomepage.htm> (la definición está disponible en): <http://www.diglib.org/about/strategic.htm>
- DLI 1. Digital Libraries Initiative 1. 1995. Digital Libraries Federation. Disponible noviembre 3, 2005 en: <http://www.dli2.nsf.gov/dlione/>
- DLI 2. Digital Libraries Initiative 2. 1998. Digital Libraries Federation. Disponible noviembre 3, 2005 en: <http://www.nsf.gov/pubs/1998/nsf9863/nsf9863.htm>
- DLTC. 1997. [Canadian] Digital Libraries Technical Committee. Sitio Web Oficial. Disponible noviembre 3, 2005 en: http://civic3.rutgers.edu/ieee_dltf.html
- Eco, Umberto. 1980. *El Nombre de la Rosa*. Trad. Ricardo Pochtar. Ed. Lumen. ISBN: 84-264-2622-0.
- eLib Programme. 2000. *Summative Evaluation of Phases 1 and 2 of the e-Lib Initiative*. JISC. Disponible noviembre 3, 2005 en: <http://www.ukoln.ac.uk/services/elib/info-projects/phase-1-and-2-evaluation/overview.pdf>
- Ercegovac, Zorana. 1999. "Introduction" within the Special Topic Issue, "Integrating Multiple Overlapping Metadata Standards". En: *Journal of the American Society for Information Science*, v. 50, n. 13, p. 1165-1168.

Esposito, Joseph. J. 2003. "The Processed Book". En: *First Monday*. Disponible noviembre 3, 2005 en: http://www.firstmonday.dk/issues/issue8_3/esposito

Esteve Botey, Francisco. 1949. "Ex-libris y ex-libristas". En: *Madrid, Coleccionismo* núm. 170 al 181, año XVI y XVII, 1929-30. Madrid, Aguilar.

European Commission. (1997). *Libro Verde Sobre la Convergencia de los Sectores de Telecomunicaciones, Medios de Comunicación y Tecnologías de La Información y Sobre sus Consecuencias para la Reglamentación*. COM 623. Disponible noviembre 3, 2005 en: <http://europa.eu.int/ISPO/convergencecp/greenp.html>

European Commission. (1997). *Propuesta para los Derechos Afines en la Sociedad de la Información [COM (97) 628 final COD97/0359*. Disponible noviembre 3, 2005 en: <http://europa.eu.int/scadplus/leg/es/lvb/l24152.htm>

European Council of Information Associations (ECIA). 1999. *Euroréférentiel I&D. "Référentiel des Compétences des Professionnels Européens de l'Information et Documentation*. Paris: ADBS Editions. ISBN 2-84365-027-5. ISSN 1263-9265. Disponible noviembre 3, 2005 en: http://www.adbs.fr/uploads/ouvrages/1917_fr.pdf

European Travel Commission. 2004. *New media Review*. Disponible noviembre 3, 2005 en: <http://www.etcnewmedia.com/review/default.asp?SectionID=10>

Fishman, Stephen. 2004. *The Copyright Handbook: How to Protect & Use Written Works*. 8th Edition. ISBN: 1-4133-0091-X. 512 p.

FOLDDOC. Free On-Line Dictionary of Computing. Disponible noviembre 3, 2005 en: <http://foldoc.doc.ic.ac.uk/foldoc/index.html>

- Foreign Policy. 2004. Página Web de la empresa. Disponible noviembre 3, 2005 en: http://www.foreignpolicy.com/story/cms.php?story_id=2594
- Foucault, Michel. 1977. *What is an author?: Language, Counter-memory, Practice: Selected Essays and Interviews*. Ithaca, NY: Cornell University Press.
- Fuentes, Carlos. 2004. *Texto completo del discurso inaugural del III Congreso Internacional de la Lengua Española*. Rosario, Argentina, Noviembre 17, 2004. Disponible noviembre 3, 2005 en: http://www.congresodelalengua3.ar/carlos_fuentes_inauguracion.htm
- Gallica. La Bibliothèque Numérique. 2004. Bibliothèque Nationale de France. Página Web de la Biblioteca. Disponible noviembre 3, 2005 en: <http://gallica.bnf.fr>
- Garduño, Roberto. 2004. "La Educación a Distancia como elemento del Desarrollo de la Sociedad del Conocimiento". En: *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, vol. 37, no. 18. Diciembre 2004. pp. 7-8. ISSN: 0187-358-X. Disponible noviembre 3, 2005 en: <http://www.ejournal.unam.mx/iibiblio/vol18-37/IBI03701.pdf>
- Gasaway, (Lolly) Laura N. 2001. *When Work Pass Into Public Domain*. University of North Carolina. Disponible noviembre 3, 2005 en: <http://www.unc.edu/~uncclng/public-d.htm>
- Gesner, Konrad, 1545. *Bibliotheca Universalis, Sive, Catalogus Omnium Scriptorum locupletissimus, in Tribus Linguis, Latina, Græca, & Hebraica ...* Zürich. Apud Christoph Froschauer.
- Gladney, Henry M., et al. 1994. Digital Library: Gross Structure and Requirements. En: *Proceedings of the First*

Annual Conference on the theory and practice of Digital Libraries.

Goldfarb, Charles F. 1996. *The Roots of SGML A Personal Recollection*. Disponible noviembre 3, 2005 en: <http://www.sgmlsource.com/history/roots.htm>

González Casanova, Pablo. 2004. *Las Nuevas Ciencias y las Humanidades: De la Academia a la Política*. Barcelona: Anthropos. 478 p. ISBN: 847658676-0.

Gorman, Michael. 1998. "What is the Internet". En: *The One Person Library*, vol. 15, No. 6 (6/98) p. 5.

Gorman, Michael. 2001. *Bibliographic Control or Chaos: an Agenda for National Bibliographic Services in the 21st Century*. Paper presented at the 67th IFLA General Conference and Council. Boston, USA, August 16-25, 2001.

Guimarães, José Augusto. 1998. *Moderno Profissional da Informação. Elementos Para sua Formação e Atuação no Mercosul Com Vistas Ao Terceiro Milênio*. Tercer Encuentro de Directores y Segundo de Docentes de las escuelas de Bibliotecología del Mercosur. Octubre 29-31, 1998. Santiago de Chile. Universidad Tecnológica Metropolitana.

Gutenberg Project. 2001. *Página Web Oficial del Proyecto*. Disponible noviembre 3, 2005 en: <http://promo.net/pg>

Haigh, Susan. 1996. *Optical Character Recognition as a Digitization Technology*. Disponible noviembre 3, 2005 en: <http://www.nlc-bnc.ca/9/1/p1-236-e.html>

Hara, Noriko y Kling, Rob. 1999. "Student's Frustrations with a Web-based Distance Education Course". En: *First Monday*, 4(12), (Diciembre 1999). Disponible noviembre 3, 2005 en: http://firstmonday.org/issues/issue4_12/hara/index.html

- Harvard Library Preservation. 2000. *Image Digitization. File Formats. Digital still Images*. Disponible noviembre 3, 2005 en: <http://preserve.harvard.edu/resources/imageformats.html>
- Harvard University. 2001. *Selection for Digitizing: A Decision-Making Matrix*. Disponible noviembre 3, 2005 en: <http://preserve.harvard.edu/bibliographies/matrix.pdf>
- Hughes, Carol A. 2000. "Information Services for Higher education. A New Competitive Space". En: *D-Lib Magazine*, (Diciembre 2000). Disponible noviembre 3, 2005 en: <http://www.dlib.org/dlib/december00/hughes/12hughes.html>
- Iannella, Renato. (Ed.) 1996. "Australian Digital Libraries Initiatives". En: *D-Lib Magazine*, (Diciembre 1996). Disponible noviembre 3, 2005 en: <http://www.dlib.org/dlib/december96/12iannella.html>
- Information Highway Advisory Council. 1997. "Preparing Canada for a Digital World - Final Report of the Information Highway Advisory Council". August 1997.
- IFLA. International Federation of Library Associations and Institutions. 1994. *Manifiesto de la IFLA/UNESCO Sobre la Biblioteca Pública 1994*. Traducido de: *UNESCO Public Library Manifesto 1994*. Disponible noviembre 3, 2005 en: <http://www.ifla.org/VII/s8/unesco/span.htm>
- IFLA. International Federation of Library Associations and Institutions. 1996. *Position Paper on Copyright in the Electronic Environment*. October 1996. Disponible noviembre 3, 2005 en: <http://www.ifla.org/V/press/pr961002.htm>
- IFLA. International Federation of Library Associations and Institutions. 1999. *Declaración de la IFLA Sobre las Bibliotecas y la Libertad Intelectual*. Disponible noviem-

bre 3, 2005 en: http://www.ifla.org/faife/policy/iflastat/iflastat_s.htm

Iglesias, Carmen. 2002. *De Historia y de Literatura como Elementos de Ficción*. Discurso al tomar el nombramiento como miembro de la Real Academia de la Lengua Española. Madrid, septiembre 30, 2002. Disponible noviembre 3, 2005 en: [http://www.rae.es/rae/gestores/gespub000001.nsf/\(voAnexos\)/arch5838FB807DBF5039C1256C7E0045FD95/\\$FILE/ciglesias.htm](http://www.rae.es/rae/gestores/gespub000001.nsf/(voAnexos)/arch5838FB807DBF5039C1256C7E0045FD95/$FILE/ciglesias.htm)

Indiana University Digital Library Program (IUDLP). 2004. Sitio Oficial *Web*. Disponible noviembre 3, 2005 en: <http://www.dlib.indiana.edu/>

INSPIRAL. 2001. [INveStigating Portals for Information Resources And Learning]. Centre for Digital Library Research [CDLR]. Disponible noviembre 3, 2005 en: <http://inspiral.cdlr.strath.ac.uk/>

Instituto de Investigaciones Jurídicas, Universidad Nacional Autónoma de México (UNAM). 2000. *Secretarías de Estado*. Gobierno de México. Disponible noviembre 3, 2005 en: <http://www.juridicas.unam.mx/navjus/gob/mx/pe1.htm>

Internet World Stats. 2004. *Usage and Population Statistics*. Disponible noviembre 3, 2005 en: <http://www.internetworldstats.com/top20.htm>

International Organization for Standardization. 2003. *The Dublin Core Metadata Element Set*. Information and Documentation ISO 15836:2003(E). Disponible noviembre 3, 2005 en: <http://www.niso.org/international/SC4/n515.pdf>

Johnson, David W., Johnson, Roger T. & Smith, Karl A. 1991. "Cooperative Learning: Increasing College Faculty Instructional Productivity", En: *ASHE-ERIC Higher Education*

- Report No. 4, 1991.* Association for the Study of Higher Education, ERIC Clearinghouse on Higher Education, Washington, D.C.
- Johnson, Mark. 1999. *XML for the Absolute Beginner*. Disponible noviembre 3, 2005 en: <http://www.cnn.com/TECH/computing/9904/02/xml.idg/>
- Joint ICSU Press / UNESCO Expert Conference on Electronic Publishing In Science. 1996. UNESCO, Paris, 19-23 February 1996. Recommendations. Disponible noviembre 3, 2005 en: <http://gateway.library.uiuc.edu/icsu/recomm~1.htm>
- Joint Pothographic Experts Group. 2003. *Página Web Oficial del Grupo de Expertos*. Disponible noviembre 3, 2005 en: <http://www.jpeg.org>
- Johnson, Doug. 2002. "Why Do We Need Libraries When We Have the Internet?". En: *Real Questions, Good Answers*, vol. 2 no. 1. Disponible noviembre 3, 2005 en: <http://www.doug-johnson.com/dougwri/internet.html>
- Kaufman, David. 1989. "Post-secondary Distance Education in Canada: Policies, Practices, and priorities". En: *R. Sweet (Ed.)*. Athabasca, Alberta: Athabasca University / Canadian Society for Studies in Education.
- Kenneth, Arnold. 1995. "The Electronic Librarian Is a Verb / The Electronic Library Is Not a Sentence". En: *Journal of Electronic Publishing*. 1:95. Disponible noviembre 3, 2005 en: <http://www.press.umich.edu/jep/works/arnold.eleclib.html>
- Kling, Rob y McKim, Geoffrey. 1997. *A Typology for Electronic Journals: Characterizing Scholarly Journals by Their Distribution Forms*. (Working Paper No. WP-97-07) Indiana University, Bloomington, Center for Social Informatics.

Disponible noviembre 3, 2005 en: <http://www.slis.indiana.edu/CSI/WP/wp97-07B.html>

Kominek, John. 1995. *Introduction to Fractal Compression*. Disponible noviembre 3, 2005 en: <http://www.faqs.org/faqs/compression-faq/part2/section-8.html>

Lafuente, Ramiro. 1999. *Biblioteca Digital y Orden Documental*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas. ISBN: 968-36-7075-X.

Lange, Holley y Winkler, B. Jean. 1997. "Taming the Internet: Metadata, a Work in Progress". En: *Advances in Librarianship*, ed. Irene Godden, v. 21, p. 47-72. San Diego: Academic Press.

Leiner, Barry M. 1998. *The Scope of the Digital Library*. Draft Prepared by Barry M. Leiner for the D-Lib Working Group on Digital Library Metrics. January 16, 1998. Revised October 15, 1998. Disponible noviembre 3, 2005 en: <http://www.dlib.org/metrics/public/papers/dig-lib-scope.html>

Leopold's Home Video Formats page. 1999. Disponible noviembre 3, 2005 en: <http://www.cs.tut.fi/~leopold/Ld/VideoFormats.html>

Lessig, Lawrence. 2001. *The Future of Ideas: The Fate of the Commons in a Connected World*. New York: Random House. ISBN 0-375-50578-4.

"The Letter of Aristeeas to Philocrates". ca. 150 b.C. En: The Forgotten Books of Eden. Chapter I, Paragraphs 11-13. Translated in the late 1800's by Dr. S. C. Malan and Dr. E. Trumpp. Translated into King James English from both the Arabic version and the Ethiopic version which was then published in The Forgotten Books of Eden in 1927 by The World Publishing Company. Disponible noviembre

- 3, 2005 en: <http://www.earth-history.com/Pseudepigrapha/FB-Eden/aristeas-intro.htm>
- Levy, David M. y Marshall, Catherine C. 1994. "Washington's White Horse? A look at Assumptions Underlying Digital Libraries". En: *Digital Libraries DL '94*. Disponible noviembre 3, 2005 en: <http://www.csd.tamu.edu/DL94/paper/levy.html>
- Library of Congress Web Site. *National Digital Library Program*. Disponible noviembre 3, 2005 en: <http://memory.loc.gov/ammem/dli2/html/lcndlp.html>
- Library of Congress Preservation Web Site. 1999. *Selection Criteria for Preservation Digital Reformatting*. Disponible noviembre 3, 2005 en: <http://lcweb.loc.gov/preserv/prd/presdig/pressselection.html>
- Licklider, Joseph C. R. 1965. *Libraries of the Future*. The MIT Press: Cambridge, MA. pp. 6-10.
- Lyman, Peter y Varian, Hal. 2000. "How Much Information?". En: *Journal of Electronic Publishing*. Diciembre, 2000, vol. 6:2. ISSN 1080-2711. Disponible noviembre 3, 2005 en: <http://www.press.umich.edu/jep/06-02/lyman.html>
- Lynch, Clifford y García-Molina, Héctor. 1995. "Interoperability, Scaling, and the Digital Libraries Research Agenda: A report on the May 18-19, 1995 IITA Digital Libraries Workshop". Disponible noviembre 3, 2005 en: <http://diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.html#2>
- Lynch, Clifford. 1997. "Searching the Internet". En: *Scientific American*. March 1997. Disponible noviembre 3, 2005 en: <http://www.hackvan.com/pub/stig/articles/trusted-systems/0397lynch.html>

- Marcum, Deanna B. 1997. "Digital Libraries: For Whom? For What?" En: *The Journal of Academic Librarianship*, March 1997, 82.
- Matthiae, Paolo. 1980. *Ebla: an Empire Rediscovered*. Trans. by Christopher Holme. Garden City, N.Y.: Doubleday. Citado por: Wellisch, Hans. 1981. "Ebla: the World's Oldest Library". En: *Journal of Library History*. vol. 16, no.3 (Summer 1981). p. 488-499.
- May, Mike. 1996. "Fractal Image Compression". En: *Science Observer*, (Sept.-Oct. 1996). Disponible en: <http://www.americanscientist.org/Issues/Sciobs96/Sciobs96-09image.html>
- Maxwell, Christine. 2000. *Global Trends that Will Impact Universal Access*. Document Submitted to UNESCO by Christine Maxwell (Ed.) on behalf of the Internet Society. July 15th, 2000. Disponible noviembre 3, 2005 en: <http://www.isoc.org/isoc/unesco-paper.shtml>
- McKiernan, Gerry. 2000. *EJI: A Registry of Innovative E-Journal Features and Functionalities*. Iowa State University. Disponible noviembre 3, 2005 en: <http://www.public.iastate.edu/~CYBERSTACKS/EJI.htm>
- McLuhan, Marshall. 1964. *Understanding Media: The Extensions of Man*. New York : McGraw-Hill.
- Messier, Paul. 2003. *Video Format Identification Guide*. Boston Art Conservation. Disponible noviembre 3, 2005 en: <http://www.video-id.com/>
- México. Academia de Ciencias. 1999. *México Frente a la Era de la Información*. México: Academia de Ciencias. 54 pp.
- México. Ley Federal del Derecho de Autor. (1996). Disponible noviembre 3, 2005 en: <http://info4.juridicas.unam.mx/ijure/tcfed/126.htm?s=>

- México. Presidencia de la República, 2001. *Plan Nacional de desarrollo 2001-2006. Capítulo Crecimiento con Calidad*. Disponible noviembre 3, 2005 en: <http://pnd.presidencia.gob.mx/index.php?idseccion=45#>
- México. Secretaría de Comunicaciones y Transportes. 2001. *Programa Sectorial de Comunicaciones y Transportes 2001-2006. Sistema Nacional e-México*. Primera Ed. ISBN 968-803-353-7. pp. 253-272.
- México. Secretaría de Comunicaciones y Transportes. 2003. *Sector Comunicaciones y Transportes; Programa 2003*. Disponible noviembre 3, 2005 en: http://www.sct.gob.mx/progratrab2003/10-P2003_e-MEXICO.pdf
- México. Secretaría de Comunicaciones y Transportes. 2003. *Sector Comunicaciones y Transportes; Programa 2004*. Disponible noviembre 3, 2005 en: http://www.sct.gob.mx/progratrab2004/c4_e-mexico_p2004.pdf
- México. Secretaría de Desarrollo Social. 2004. *Gobierno en-línea*. Disponible noviembre 3, 2005 en: <http://www.sedesol.gob.mx/gobiernoenlinea/emexico.htm>
- México. Secretaría de Educación Pública. 2001. *Programa Nacional de Educación 2001-2006*. Primera Ed. ISBN: 970-18-6213-9. Disponible noviembre 3, 2005 en: http://www.sep.gob.mx/wb2/sep/sep_2734_programa_nacional_de
- Millares Carlo, Agustín. 1971. *Introducción a la Historia del Libro y de las Bibliotecas*, México: Fondo de Cultura Económica.
- Mintzer, Fred. 1999. Long Term Storage of Digital Media Objects for Digital Libraries, En: *NSF Workshop on Data Archival and Information Preservation*, Washington DC, March 1999.

- MODS. Metadata Object Description Schema. 2004. Página Web oficial. Washington: Library of Congress, Disponible noviembre 3, 2005 en: <http://www.loc.gov/standards/mods/>
- Mogge, Dru. 1999. "Seven Years of Tracking Electronic Publishing: The ARL Directory of Electronic Journals, Newsletters and Academic Discussion Lists". En: *Library Hi Tech*. 17:1, (1999). pp. 17-25
- Molina, M^a Clemencia y Pérez, Marta. 1988. "El Currículo y los Perfiles del Profesional en Bibliotecología y Ciencias de la Información". En: *Revista Interamericana de Bibliotecología*. Medellín: Colombia. 11(1).
- Moore, Michael G. y Kearsley, Greg. 1996. *Distance Education: a System's View*. Belmont, Ca.: Wadsworth Pub. Co.
- Morales C., Estela. 2001. "La Sociedad de la Información en el Siglo XXI y la Biblioteca Universitaria". En: *Revista Digital Universitaria*. Vol 2, no. 2. Junio 30 del 2001. México: Dirección General de Servicios de Cómputo Académico, Universidad Nacional Autónoma de México, (UNAM). Disponible noviembre 3, 2005 en: <http://www.revista.unam.mx/vol.2/num2/art1/>
- Morales C., Estela, 2004. "El Hombre y la Sociedad como Centro de los Servicios de Información". En: *XXII Coloquio de Investigación Bibliotecológica y de la Información: Sociedad, Información y Bibliotecas*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas. Septiembre 22-24, 2004.
- Morrison, Alan et al. 1998. "Creating and Documenting Electronic Texts. Chapter 3: Digitization — Scanning, OCR, and Re-keying". En: *AHDS Guides to Good Practice*. ISSN: 1463-5194. Disponible noviembre 3, 2005 en: <http://ota.ahds.ac.uk/documents/creating/chap3.html#1>

- NCOIT&R. National Coordination Office for Information Technology Research & Development. (1994). Disponible noviembre 3, 2005 en: <http://www.hpcc.gov> y en: <http://www.hpcc.gov/about/presentations/furlani-nhpcc-3apr02/>
- Nationmaster. 2004. Massive Central Data Source. Página Web de la empresa. Disponible noviembre 3, 2005 en: <http://www.natinmaster.com>. Véase también: Personal computers (top 100 countries) en: http://www.nationmaster.com/graph-T/med_per_com e Internet users per capita (top 100 countries) en: http://www.nationmaster.com/graph-T/med_per_com_cap
- NII. National Information Infrastructure. (1998). *Relationship Between HPCC and NII*. Disponible noviembre 3, 2005 en: <http://www.itrd.gov/pubs/iita/1.1.html>
- (NSDL) National Science, Technology, Engineering, and Mathematics Education Digital Library. 2001. Página Web de la Biblioteca. Disponible noviembre 3, 2005 en: <http://nsdl.org>
- Naur, Peter. (ed). 1963. *Revised Report on the Algorithmic Language Algol 60*. Communications ACM 6:1. pp 1-17.
- Nelson, (Ted) Theodor H. 1974. *Computer Lib / Dream Machines*. Ed. orig. Mindful Press, 1974. 153 pp. ISBN: 0893470023. Ed. revisada. Redmond, WA.:Tempus Books of Microsoft Press, 1987. ISBN: 0-914-84549-7.
- Nipper, Søren. 1989. "Third Generation Distance Learning and Computer Conferencing". En: *Mason, R. & Kaye, A. (eds). Mindweave: Communications, Computers and Distance Education*. Capítulo 5, pp. 63-73. Oxford: Pergamon.
- Núñez, Andrés. 2000. *Una comparación del campus virtual de British Open University y el campus virtual de Florida State University: constructivismo vs. Conductismo*.

Centro Virtual Cervantes. Instituto Cervantes (España). Disponible noviembre 3, 2005 en: http://cvc.cervantes.es/obref/formacion_virtual/campus_virtual/nunez.htm

Núñez, Andrés. 2001. "La Educación Digital". En: *Congreso Internacional de Tecnología, Educación y Desarrollo Sostenible*. Murcia, España. 17-19 septiembre 2001. Disponible noviembre 3, 2005 en: <http://www.edutec.es/edutec01/edutec/comunic/TSE16.html>

Nürnberg, P.J., Furata, R., Legget, J.J., Marshall, C.C. y Shipman, F.M. 1995. "Digital Libraries: Issues and Architecture". En: *The Second Annual Conference on Theory and Practice of Digital Libraries* (Texas A&M University, College Station, Texas, (Junio 1995), pp. 148-154. Disponible noviembre 3, 2005 en: <http://www.csdl.tamu.edu/DL95/papers/nuernberg/nuernberg.html>

Oasis Standards and Other Approved Work. 2005. Página Oficial de la Organización OASIS. Disponible noviembre 3, 2005 en: <http://www.oasis-open.org/specs/index.php#opendocumentv1.0>

ODLIS (On-Line Dictionary of Library and Information Science). 2002. Western Connecticut State University. Disponible noviembre 3, 2005 en: <http://www.wcsu.edu/library/odlis.html>

O'Donnell, James J. 1994. "The Virtual Library: An Idea Whose Time Has Passed". En: *Okerson and Mogge, eds., Gateways, Gatekeepers, and Roles in the Information Omniverse*. Washington, D.C.: Association of Research Libraries, 1994, pp. 2-5, 19-31. Disponible noviembre 3, 2005 en: <http://ccat.sas.upenn.edu/jod/virtual.html>

Olson, David R., y Bruner, Jerome. 1974. "Learning Through Experience and Learning Through Media" En: *Media and Symbols: The Forms of Expression*. Chicago: The University of Chicago Press. (The 73rd NSSE Yearbook).

- Open eBook Forum. 2002. *Open eBook Publication Structure Specification Version 1.2*. Disponible noviembre 3, 2005 en: <http://www.openebook.org/oebps/oebps1.2/index.htm>
- Organización de las Naciones Unidas. Programa de las Naciones Unidas para el Desarrollo (PNUD). 2001. *Informe Sobre Desarrollo Humano 2001: Poner el Adelanto Tecnológico al Servicio del Desarrollo Humano*. Disponible noviembre 3, 2005 en: <http://www.undp.org/bdr2001/spanish/>
- Otlet, Paul M. 1934. *Traité de Documentation: Le Livre Sur le Livre*. Bruxelles: Editions Mundaneum Palais Mondial. Reimpreso por: Liège: Centre de Lecture Publique de la Communauté Française en Belgique, 1989. Reimpreso en español como: Otlet, Paul. 1996. *El Tratado de Documentación: El Libro Sobre el Libro. Teoría y Práctica*. (Traducción de M^a Dolores Ayuso). Murcia: Universidad de Murcia, Caja de Ahorros de Murcia. ISBN: 84-7684-766-1.
- Pérez Mazatán, Javier. 2005. "e-México National System". En: *Fourth India Mexico Joint Commission*. Nueva Delhi, India: Octubre 21-22, 2005. Presentación *Power Point*. Láminas 32-35. Disponible noviembre 3, 2005 en: http://www.e-mexico.gob.mx/wb2/eMex/eMex_the_emexico_nationa
- Planning Workshop on Distributed Knowledge Work Environments. 1997. *Final Report of the Planning Workshop: Digital Libraries*. [Introducción]. Disponible noviembre 3, 2005 en: <http://www.si.umich.edu/SantaFe>
- Popper, Karl R. 1983. *Conjeturas y Refutaciones: El Desarrollo del Conocimiento Científico*. Buenos Aires: Paidós. p. 280.

- Public Communication Group, 2003. "Trends in Electronic Journal Collection: An International Perspective". Summary. Joshua Clarke, (Ed.). En: *PCG Vantage*, vol. 1:2, November 2003. Disponible noviembre 3, 2005 en: http://www.imakenews.com/pcgplus/e_article000200562.cfm
- Rader, Hannelore B. 1997. "Educating Students for the Information Age: The Role of the Librarian". En: *Reference Services Review*, 25:2, (Summer 1997).
- Raitt, David. 2000. "Digital Libraries Initiatives Across Europe". En: *Computers in Libraries*. 20:10, (Nov/Dic. 2000). Disponible noviembre 3, 2005 en: <http://www.infotoday.com/cilmag/nov00/raitt.htm>
- Rangel, David. 1998. *Derecho Intelectual*. México: McGraw-Hill, p. 111.
- Rice, Condoleezza. 1998. *Copyright & Fair Use. Memorandum to Faculty Members*. Stanford University Libraries. Disponible noviembre 3, 2005 en: http://fairuse.stanford.edu/stanford_notices/rice.html
- Roes, Hans. 2001. "Digital Libraries and Education". En: *D-Lib Magazine*. 7:7/8, (Julio/Agosto 2001). Disponible noviembre 3, 2005 en: <http://www.dlib.org/dlib/july01/roes/07roes.html>
- Roush, Wade. 2001. "e-Book Basics: e-Book Primer". En: *e-Book Web*. Disponible noviembre 3, 2005 en: <http://www.ebookweb.org/basics/ebook.primer.htm>
- Rusch-Feja, Diann, y Siebeky, Uta. 1999. "Evaluation of Usage and Acceptance of Electronic Journals". En: *D-Lib Magazine*. 5:10. (1995). Disponible noviembre 3, 2005 en: <http://www.dlib.org/dlib/october99/rusch-feja/10rusch-feja-full-report.html>

- Sakaiya, Taichi. 1995. *Historia del Futuro. La Sociedad del Conocimiento*. Santiago de Chile: Ed. Andrés Bello, 1995, 225 p.
- Sánchez, Alfredo. 1999. *Programa Nacional de Bibliotecas Digitales*. México: Conacyt; Red de Desarrollo de Investigación en Informática (REDII). Disponible noviembre 3, 2005 en: <http://ict.udlap.mx/dl/dlmex/pnbd-version3.html>
- Sayood, Khalid. 2000. *Introduction to Data Compression*. 2nd ed. San Francisco : Morgan Kaufmann Publ. ISBN: 1-55 860-558-4. pp. 1-37.
- Santa Fe Planning Workshop on Distributed Knowledge Work Environments. 1997. Final Report on Digital Libraries. March 9-11, 1997. Report Version: September 20, 1997. Disponible noviembre 3, 2005 en: <http://www.si.umich.edu/SantaFe/>
- Seamans, Nan y Mcmillan, Gail. 1998. *Digital Library Definition for DLI2*. Disponible noviembre 3, 2005 en: <http://scholar.lib.vt.edu/DLI2/defineDL.html>
- Senso, José A. y Rosa Piñero, Antonio de la. 2003. "El Concepto de Metadato: Algo Más que Descripción de Recursos Electrónicos". En: *Ciência da Informação. (Ci. Inf.)*, mayo-agosto 2003, vol.32, no.2, pp.95-106. ISSN: 0100-1965.
- Serrano S., Arturo y Martínez M., Evelio. 2003. *La Brecha Digital : Mitos y Realidades*. Mexicali: Universidad Autónoma de Baja California, FOECA (Fondo Editorial de Baja California) 175 p. ISBN: 970-9051-89-X.
- Shannon, Claude. 1948. "A Mathematical Theory of Communication". En: *The Bell System Technical Journal. Bell Laboratories*, (1948). pp. 379-636.
- Sharon, Taly y Frank, Ariel J. 2000. "Digital Libraries on the Internet". En: *66th IFLA Council and General Conference*.

- Jerusalem, 13-18 August 2000. Disponible noviembre 3, 2005 en: <http://www.ifla.org/IV/ifla66/papers/029-142e.htm>
- Shera, Jesse H. 1990. "Lo que el Bibliotecario Necesita Saber". En: *Los Fundamentos de la Educación Bibliotecológica*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas. pp. 201-231.
- Sistema Nacional e-México. (2001). Disponible noviembre 3, 2005 en: <http://www.e-mexico.gob.mx>
- Smith, Abby. 1999. *Why Digitize?* CLIR. Council on Library and Information Resources. Commission on Preservation and Access. 2:99, ISBN 1-887334-65-3. Disponible noviembre 3, 2005 en: <http://www.clir.org/pubs/reports/pub80-smith/pub80.html>
- Soules, Aline. 2002. "The Rights and Wrongs of the Doctrine of the First Sale". En: *Electronic Book Web*. Disponible noviembre 3, 2005 en: [http://12.108.175.91/ebookweb/stories/storyReader\\$1816](http://12.108.175.91/ebookweb/stories/storyReader$1816)
- Special Libraries Association (SLA). 1996. *Competencies for Special Librarians of the 21st Century. Full Report*. Submitted to the SLA Board of Directors by the Special Committee on Competencies for Special Librarians. May 1996. Revised June 2003. Disponible noviembre 3, 2005 en: <http://www.sla.org/content/SLA/professional/meaning/competency.cfm>
- Spink, Amanda. 1995. "Digital Libraries and Sustainable Development?" En: *Digital Libraries DL '95*. Disponible noviembre 3, 2005 en: <http://www.cSDL.tamu.edu/DL95/papers/spink/spink.html>
- Sterling, J. A. L. 2000. *World Copyright Law*. London: Sweet & Maxwell. 1084 pp.

- Stoker, David. 1999. "Wanted — An Innovative and Visionary Evidence Based/Knowledge Management Librarian." En: *Journal of Librarianship and Information Science*. 31(2): 67-69.
- Teasdale, Guy. 1995. *L'Hypertexte: Historique et Applications en Bibliothéconomie*. Disponible noviembre 3, 2005 en: <http://www.ebsi.umontreal.ca/cursus/vol1no1/teasdale.html>
- Tennant, Roy. 1999. "Digital v. Electronic v. Virtual Libraries". En: *Berkeley Digital Library SunSITE*. Disponible noviembre 3, 2005 en: <http://sunsite.berkeley.edu/mydefinitions.html>
- Tenopir, Carol. 2003. *Electronic Publishing: Research Issues for Academic Librarians and Users*. Disponible noviembre 3, 2005 en: http://www.findarticles.com/p/articles/mi_m1387/is_4_51/ai_105046543/pg_1
- Thompson, James W. 1967. *The Medieval Library*. New York: Hafner Publishing.
- Thompson, Kim. 1998. *University of California Selection Criteria for Digitization*. Disponible noviembre 3, 2005 en: <http://www.library.ucsb.edu/ucpag/digselec.html>
- Tichenor, Philip J., Donohue, George y Olien, Clarice. 1970. "Mass Media Flow and Differential Growth in Knowledge". En: *Public Opinion Quarterly*, 34, pp. 159-170.
- Torre Villar, Ernesto de la. 1994. *Ex Libris y Marcas de Fuego*. México: UNAM, Dirección General de Publicaciones y Fomento Editorial. ISBN: 968-36-7708-8.
- Torres, Araceli. 2000. *La Biblioteca Virtual : ¿Qué es y qué promete?* México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas. ISBN: 968-36-8927-2.

- Tramullas, Jesús. 2002. "Las Bibliotecas Como Sistemas Multifacetados de Información". Sesión en el programa de doctorado "Informática y Modelización Matemática", Universidad Rey Juan Carlos, España. Disponible noviembre 3, 2005 en: <http://www.tramullas.com/presentaciones.html>
- Turner, Frederick y Corbacho, Alejandro. 2000. *Las Nuevas Funciones del Estado*. Disponible noviembre 3, 2005 en: <http://www.unesco.org/issj/rics163/abstracts163.htm>
- Twigg, Carol y Miloff, Michael. 1998. "The Global Learning Infrastructure". En: *Don Tapscott, Alex Lowry and David Ticoll, Blueprint to the Digital Economy*, McGraw-Hill, 1998.
- UCLA-NSF Social Aspects of Digital Libraries Workshop. 1996. *Final Report. UCLA-NSF*. Febrero 15-17 1996. Disponible noviembre 3, 2005 en: <http://is.gseis.ucla.edu/research/dl/>
- UNESCO. 2002. *Declaración Universal de la UNESCO Sobre la Diversidad Cultural*. Página Web de la UNESCO. Disponible noviembre 3, 2005 en: http://www.unesco.org/culture/pluralism/diversity/html_sp/index_sp.shtml
- United States of America. House of Congress. 1996. *Fair Use Guidelines for Educational Multimedia established by Congress of the United States: Nonlegislative Report of the Subcommittee on Courts and Intellectual Property Committee on the Judiciary US House of Representatives*. Disponible noviembre 3, 2005 en: <http://www.libraries.psu.edu/mtss/fairuse/guidelines.html>
- United States of America. House of Congress. 1998. *Digital Millennium Copyright Act. Executive Summary*. Disponible noviembre 3, 2005 en: http://www.copyright.gov/reports/studies/dmca/dmca_executive.html
- University of Cornell. Electronic Publishing Steering Committee. 1998. *Electronic Publishing Strategies for Cornell*

- University. Disponible noviembre 3, 2005 en: <http://www.library.cornell.edu/ulib/pubs/EPSCFinalReport1998.htm>
- University of Maryland: University College. 2000. *Copyright and Fair Use in the Classroom, on the Internet, and the World Wide Web*. Disponible noviembre 3, 2005 en: <http://www.umuc.edu/library/copy.html>
- University of Texas System Digital Library. 2002. "TILT: [Texas Information Literacy Tutorial]". Disponible noviembre 3, 2005 en: <http://tilt.lib.utsystem.edu>
- UWE [University of Western England]. 2002. *Networked Learning Supported Framework*. Disponible noviembre 3, 2005 en: <http://www.uwe.ac.uk/library/nlsf/>
- Verduin, John R. y Clark, Thomas A. 1991. *Distance Education: The Foundations of Effective Practice*. San Francisco: Jossey-Bass. pp. 16-42 y 81-83.
- Villate, Javier. 2001, *La Propiedad Intelectual en la Nueva Era Digital*. Archivo del Observatorio para la CiberSociedad. Disponible noviembre 3, 2005 en: <http://www.cibersociedad.net/archivo/articulo.php?art=40>
- Voutssás M., Juan. 1995. "Crónica de Tres Muertes Anunciadas: El Teatro, la Radio y el Libro". En: *BIBLIOS: Revista para Especialistas en Información*, vol. I, no. 2, pp. 2-3. marzo-abril 1995.
- Watstein, Sarah, Calarco, Pascal y Ghaphery, James. (1999) "Digital Libraries: Keywords", En: *RSR/Reference Services Review*, 27 (4), pp. 344-352.
- Wattenberg, Frank. 1998. "A National Digital Library for Science, Mathematics, Engineering and Technology Education", En: *D-Lib Magazine*, (Octubre 1998). Disponible noviembre 3, 2005 en: <http://www.dlib.org/dlib/october98/wattenberg/10wattenberg.html>

Weber, Hartmut y Dörr, Marianne. 1997. *Digitization as a Method of Preservation?* CLIR. Council on Library and Information Resources. Commission on Preservation and Access. 10:97. ISBN 1-887334-53-6. Disponible noviembre 3, 2005 en: <http://www.clir.org/pubs/reports/digpres/digpres.html>

What is MARC. 1996. Disponible noviembre 3, 2005 en: <http://www.ifla.org/VI/3/p1996-1/unimarc.htm>

World Economic Forum. 2004. *Global Information Technology Report 2003-2004*. Oxford: Oxford University Press. ISBN: 0-19-517361-9.

World Summit on the Information Society (WSIS). 2003. Cumbre Mundial sobre la Sociedad de la Información. *Declaración de Principios*. Ginebra, Diciembre 10-12, 2003. Disponible noviembre 3, 2005 en: http://www.itu.int/dms_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0004!!pdf-S.pdf

World Wide Web Consortium [W3C]. 2002. *XML*. Disponible noviembre 3, 2005 en: <http://www.w3.org/xml>

World Wide Web Consortium [W3C]. 2002. *PNG*. Disponible noviembre 3, 2005 en: <http://www.w3.org/Graphics/PNG>

Writing HTML. 1998. Disponible noviembre 3, 2005 en: http://www.mcli.dist.maricopa.edu/tut_es/

Wyner, Aaron D. *The Significance of Shannon's Work*. Disponible noviembre 3, 2005 en: <http://cm.bell-labs.com/cm/ms/what/shannonday/work.html>

The XML Cover Pages - Home page. 2001. Disponible noviembre 3, 2005 en: <http://oasis-open.org/cover/>

Zeng, L. y Zhang, J. 2002. "Digital Libraries; Where to Go?: An Analysis of Definitions, Architectures and Projects of

Digital Libraries". En: *Journal of the China Society for Scientific & Technical Information*. 19:1, February. p. 64-73.

Zia, Lee L. 2000. "The NSF National Science, Mathematics, Engineering and Technology Education Digital Library Program". En: *D-Lib Magazine*, (October 2000). Disponible noviembre 3, 2005 en: <http://www.dlib.org/dlib/october00/zia/10zia.html>

Ziv, Jacob y Lempel, Abraham. 1977. "A Universal Algorithm for Sequential Data Compression". En: *IEEE Transactions on Information Theory*, IT-23:3. (Mayo 1977).

OTROS SITIOS DE REFERENCIA DE ESPECIAL INTERÉS ACERCA DE BIBLIOTECAS DIGITALES:

D-Lib Magazine (Digital Library Magazine). Copyright© 2005 Corporation for National Research Initiative (CNRI). Mensual. ISSN 1082-9873. Sitio Web De la Revista. Disponible 3 de marzo, 2005 en: <http://www.dlib.org>

The Journal of Electronic Publishing. Published by the Columbia University Press. ISSN 1080-2711. Cuatrimestral. Sitio Web De la Revista. Disponible 3 de marzo, 2005 en: <http://www.press.umich.edu/jep>

Library Literature & Information Science Full Text. © 2005 The HW Wilson Company. Disponible noviembre 3, 2005 en: <http://www.hwwilson.com/Databases/liblit.htm>

Libweb: Library Servers Via www. Sitio Web. Actualizado todos los días, a la medianoche. Tiempo del pacífico. Libweb contiene listas de más de 7200 páginas Web de bibliotecas en más de 125 países. Disponible noviembre 3, 2005 en: <http://sunsite.berkeley.edu/Libweb/>

SWIFT. Selected Digital Library Internet Resources. Columbia University. Sitio Web de la Universidad de Columbia. Dis-

ponible noviembre 3, 2005 en: <http://www.columbia.edu/cu/libraries/inside/projects/diglib/resources.html>

Librarians' Index to the Internet (lii.org). Una guía por materias construida por bibliotecarios a 6,000 sitios de recursos en la Internet. Disponible noviembre 3, 2005 en: <http://sunsite.berkeley.edu/Help/searchdetails.html>

International Federation of Library Associations (IFLA). Digital Libraries: Resources and Projects. Página Web de la IFLA. Disponible noviembre 3, 2003 en: <http://www.ifla.org/II/diglib.htm>

GT. Bibliotecas Virtuais. "Exemplos de Projetos de Biblioteca Virtual no Mundo". Disponible noviembre 3, 2003 en: <http://www.cg.org.br/gt/gtbv/mundo.htm> Digital Libraries Federation. Página Web Oficial. Disponible noviembre 3, 2005 en: <http://www.diglib.org/sitemap.htm>

Preservation: A Selected Bibliography. Amigos Library Services. Una página con una extensa bibliografía acerca de preservación, incluida la digital. Página Web. Disponible noviembre 3, 2005 en: <http://www.amigos.org/preservation/bibpress.html>

CLIR Reports y CLIR Issues. The Council of Library and Information Resources. Reportes profesionales de numerosas investigaciones acerca de variados temas en bibliotecas digitales. Página Web. Disponible noviembre 3, 2005 en: <http://www.clir.org/pubs/reports/reports.html> y en: <http://www.clir.org/pubs/issues/issues.html>

Bibliothèques Numériques. Portal de cultura francesa con ligas a numerosas bibliotecas digitales del mundo franco-parlante. Disponible noviembre 3, 2005 en: <http://www.culture.fr/BibliothequesMediatheques/c391>

II Congreso ONLINE del Observatorio para la CiberSociedad: ¿Hacia qué sociedad del conocimiento? Barcelona, 30

oct. al 16 nov. del 2004. Disponible noviembre 3 del 2005 en: http://www.cibersociedad.net/congres2004/index_es.html

Bibliotecas del Mundo. En la Biblioteca Cervantes Virtual. Directorio de enlaces a bibliotecas virtuales, catálogos en línea, proyectos de digitalización, hemerotecas, Webs e instituciones relacionadas con la literatura y la cultura hispanas. Disponible noviembre 3, 2005 en: http://www.cervantesvirtual.com/bib_mundo/biblioteca_mundo.shtml?portal=2

BASES DE DATOS DOCUMENTALES SOBRE CONSERVACIÓN

Bibliographic Database of the Conservation Information Network (BCIN) (ca. 200,000 citas). Disponible noviembre 3, 2005 en: http://www.bcin.ca/English/home_english.html

Abstracts of International Conservation Literature (AATA Online) (100,000+ abstracts). Disponible noviembre 3, 2005 en: <http://aata.getty.edu/NPS/>

Anexos

(Todas las referencias electrónicas han sido verificadas
como existentes y exactas hasta el 3 de noviembre del 2005)

ANEXO 1:

Producción mundial estimada de contenidos originales, almacenados digitalmente, usando técnicas estándar de compresión, estimado en terabytes, para 1999. Fuente: Lyman y Varian, 2000.

Medio de almacenamiento	Tipo de contenido	Terabytes por año, estimado superior	Terabytes por año, estimado inferior	Tasa de crecimiento, %
Papel	Libros	8	1	2
	periódicos	25	2	-2
	Revistas	12	1	2
	Documentos de oficina	195	19	2
	Subtotal:	240	23	2
Película	Fotografías	410,000	41,000	5
	Cine	16	16	3
	Rayos-X	17,200	17,200	2
	Subtotal:	427,216	58,216	4
Óptico	CDs musicales	58	6	3
	CDs de datos	3	3	2
	DVDs	22	22	100
	Subtotal:	83	31	70
Magnético	Cintas de Camcorder	300,000	300,000	5
	Discos de PCs	766,000	7,660	100
	Servidores pequeños	460,000	161,000	100
	Servidores grandes	167,000	109,000	100
	Subtotal:	1,693,000	635,660	55
TOTAL:		2,120,539	693,930	50

1 Terabyte = 1,000 Gigabytes = 1,000,000 Megabytes = 1,000,000,000 Kilobytes
= 1,000,000,000,000 bytes = un trillón de bytes.

[LYMAN Y VARIAN, 2000]

ANEXO 2:

Proceso evolutivo de Internet en la pyme. Fuente: <http://www.anetcom.es/download/download.asp?seccion=Nots&FILE=InternetyPymes.doc>.

Fases	Agentes del cambio	Problemas detectados	Soluciones
FASE 1. Descubrimiento de Internet OBJETIVO: Obtener beneficios de la información presente en Internet. REQUISITO: En la empresa debe haber una persona que lidere la incorporación de las TIC.			
	Informático Técnico	Desconocimiento de los sitios web / útiles para cada Departamento	Formación personalizada a cada persona de la empresa
	Directivos y mandos intermedios	Desconocimiento de los sitios web útiles para la empresa	Formación personalizada a los directivos y mandos intermedios
FASE 2. Introducción de Internet en sus procesos internos (e-Business) OBJETIVO: Obtener beneficios por digitalizar la empresa para gestionar mejor toda la información que maneja, de una manera eficiente, segura y desde distintos lugares. REQUISITO: La empresa debe ver claro dónde están los retornos de la inversión.			
INTRANET: Con empleados y socios	Informático / Técnico en la pyme	Sin formación en gestión del cambio: Han de incentivar a diario el cambio las carreras	Formación en último curso de TODAS las facultades universitarias de todas
	Subcontratación TIC-Empresas de informática	Priorizan vender sus productos/ servicios en lugar de asegurar el retorno de inversión	Buscar asesoría imparcial
	Marketing / Ventas	Pasividad: lo ven como amenaza por retribución por objetivos	Incentivar el uso del Mercadeo por Internet para tener una retribución mayor
	Directivos y mandos intermedios	Priorizan otras actividades (ERP, Calidad, Medio ambiente, Mercadeo tradicional, etc.). Desconocen las ventajas del teletrabajo.	Priorizar la implantación de las TIC en todos los procesos. Suministrar una PC portátil y conexión de banda ancha desde el domicilio.
	TODA la plantilla	Resistencia al cambio	Motivar el uso de las TIC: Mostrar la utilidad para cada persona.

Fases	Agentes del cambio	Problemas detectados	Soluciones
FASE 3. Introducción de Internet en sus procesos externos (portales) OBJETIVO: Obtener beneficios por relacionarse con otras empresas mediante Internet. REQUISITO: La empresa debe experimentar nuevas formas de colaboración.			
EXTRANET: Con proveedores y clientes actuales	Informáticos / Técnicos de la empresa y de los clientes y proveedores actuales	Diferentes niveles de uso de las TIC en el resto de clientes y proveedores actuales	Acelerar el paso de clientes y proveedores actuales por las Fases 1 y 2
INTERNET: Con proveedores y clientes nuevos	Asociación empresarial/ Centro tecnológico	Falta agente del cambio de la asociación y de sus asociados (Fase 1)	Imitar o colaborar con otras asociaciones que estén teniendo éxito demostrado

ANEXO 3:

México – Numeralia

1. Población total: 105.9 millones de habitantes a fines del 2004
2. Población aproximada de mexicanos en EUA: 20 millones
3. Población de 15 años y más: 62,842, 638
4. Población de 15 años y más alfabetizada: 56,841,673 (90.45%)
5. Número de viviendas: 21,858,085
6. Porcentaje de viviendas que cuenta con televisión: 85.6% (18,710,521)
7. Porcentaje de viviendas que cuenta con televisión por cable: 8.9% (1,945,370)
8. Porcentaje de viviendas que cuenta con videocasetera: 38.6% (8,437,221)
9. Porcentaje de viviendas que cuenta con teléfono: 36.2% (7,912,627)
10. Porcentaje de viviendas que cuenta con computadora: 9.3% (2,032,802)
11. Porcentaje de viviendas con conexión a Internet: 4.9% (1,071,046)
12. Número de líneas telefónicas en servicio, por cada 100 habitantes: EUA 66.5, Alemania 63.5, Países Bajos 62.1, Japón 59.7, Puerto Rico 33.6, Uruguay 28.3, Chile 23.9, Brasil 21.7, Argentina 21.6, México 15.8.
13. Número de computadoras personales por cada 100 habitantes: EUA 29.9, Japón 31.5, Australia 45.2, Singapur 41.3, España 14.0, Brasil 4.2, Argentina 6.9, México 5.4.
14. Usuarios de Internet por cada 100 habitantes: Suecia 74.6, Corea del Sur 61.2, EUA 69.0, Japón 61.3, Hong Kong 70.0, CEE 29.9, Rusia 17.3, Chile, 30.1, Brasil 11.1, Uruguay 34.7, Argentina 12.3, México 13.22.
15. PC conectadas a Internet (%): EUA 50, Países Bajos 25.8, Nueva Zelanda 25, Canadá 19.7, Australia 18.2, Argentina 14.2, Brasil 11.7, México 11.2, Uruguay 7.7, Chile 5.8.
16. Costo por 20 horas de acceso a Internet (dólares): Suecia 2.6, Italia 32.2, México 24.1.

17. Total de gastos en TIC's (millones de dólares): EUA 812,635; Japón 413,772; Alemania 154,645; Reino Unido 137,726; Francia 120,569; Brasil 50,031; México 19,211; Argentina 11,642; Colombia 10,434; Chile 5,715
18. Gastos en TIC's como % del PIB: Nueva Zelanda 14.4; Suecia 11.3; Australia 10.7; Singapur 9.9; Reino Unido 9.7; Colombia 12; Brasil 8.3; Chile 8.1; Argentina 4; Venezuela 4; México 5.4.
19. Gastos en TIC's <i>per capita</i> (dólares): Japón 3256; Estados Unidos 2924; Suecia 2804; Países Bajos 2327; Reino Unido 2319; Chile 371; Argentina 310; Brasil 287; Colombia 231; Venezuela 199; México 196.
20. Índice de capital social: Finlandia 6.7; Noruega 6.6; Suiza 6.6; República Eslovaca 6.5; Dinamarca 6.5; Uruguay 5.3; Argentina 5.1; Costa Rica 5; Panamá 4.7; Chile 4.6; México 4.3.
21. Índice de competencia de las TIC en los negocios y su entorno económico: Singapur 6.2; Finlandia 6.2; Reino Unido 6.1; Islandia 6.1; Suiza 6; EUA 6; Chile 4.5; Uruguay 4.3; Costa Rica 4.1; Argentina 4; Panamá 3.8; El Salvador 3.7; México 3.4
<p>Fuentes:</p> <p>Secretaría de Gobernación. Consejo Nacional de Población y Vivienda (CONAPO). Diciembre de 2004</p> <p>2 <i>United States Census 2000</i>. US Census Bureau. (http://www.census.gov/main/www/cen2000.html).</p> <p>3, 4, 5, 6, 8, 9, 10 <i>XII Censo General de Población y Vivienda, 2000</i>. Tabulados de la Muestra Censal. Cuestionario Ampliado. México, 2000. (http://www.inegi.gob.mx)</p> <p>7, INEGI, <i>Encuesta Nacional de Ingreso y Gasto de los Hogares, 2000</i>.</p> <p>11, INEGI, <i>ENCO, Módulo Nacional de Computación Año 2001</i>.</p> <p>12, International Telecommunications Union, (ITU), 2001.</p> <p>13, 14 European Travel Commission. New Media Review. 2004, y Nationmaster.com. 2004.</p> <p>15, 16 Harvard University. <i>The Global Information Technology Report 2001-2002</i>. (http://www.cid.harvard.edu).</p> <p>17, 18, 19 The World Bank. (http://www.worldbank.org/data/countrydata/countrydata).</p> <p>20 El índice de capital social está considerado dentro del Índice de Networked Society, definido como la evaluación en la calidad del aprendizaje usando las TIC, la extensión de sus usos en el proceso de aprendizaje, la extensión de oportunidades en la industria de las TIC y factores sociales y demográficos. Harvard University. <i>The Global Information Technology Report 2001-2002: Readiness for the Networked World</i>.</p> <p>21 El Índice de competencia de las TIC en los negocios y su entorno económico está considerado dentro del índice de Networked Policy, definido como el entorno político de la información y las comunicaciones así como los negocios y el clima económico. Harvard University. <i>The Global Information Technology Report 2001-2002: Readiness for the Networked World</i>.</p>

ANEXO 4:

Tablas de tipos de datos y formatos utilizados en proyectos de digitalización.

Basada en: *"Introduction to Metadata: Pathways to Information"*. Getty Research Institute.

Página Web del Instituto. Disponible 3 de enero, 2005 en: http://www.getty.edu/research/conducting_research/standards/intrometadata/

Tipo de Datos	Aplicaciones	Formatos	Referencias
Datos Alfanúmericos	Textos de tipo literario y lingüístico para lectura, análisis, edición e intranets.	TXT, XML y PDF	<i>"Text Encoding Initiative: (TEI)"</i> . http://www.tei-c.org/ <i>"Creating and Documenting Electronic Texts"</i> . (OTA, 1999): http://ota.ahds.ac.uk/documents/creating/ and <i>"TEI Text Encoding in Libraries Guidelines for Best Practices"</i> . (DLF, 1999). http://diglib.org/standards/tei.htm
Datos Alfanuméricos	Textos literarios para distribución en la red	HTML, XML, OEB	Usar DTD's o esquema de Tipo de Elemento. http://w3c.org <i>"Open eBook Publication Structure Specification Version 1.2"</i> . http://www.openebook.org/oebps/oebps1.2/index.htm
Datos Alfanuméricos	Archivos planos, estadísticas, bases de datos jerárquicas o relacionales.	Texto ASCII o UTF-8, PDF, SPSS, SAS con metadatos para ver columnas, renglones, etc.	Igual que en el primer renglón, más <i>"Guide for Social Sciences Data Preparation & Archiving"</i> (ICPSR, 2002) http://www.icpsr.umich.edu/access/dpm.html

Tipo de Datos	Aplicaciones	Formatos	Referencias
Datos tipo imagen de documentos de textos, b/n, grises, o color	Digitalizaciones retrospectivas de libros o revistas	TIFF no comprimido o jpeg2000 sin pérdida con paleta y resolución adecuadas al tipo de material.	<p><i>"Image Digitization: File Formats, Digital Still Images"</i>. (Harvard, 2000) http://preserve.harvard.edu/resources/imageformats.html</p> <p><i>"Benchmark for Faithful Digital Reproductions of Monographs and Serials. Version 1"</i>. (DLF, 2002) http://www.diglib.org/standards/bmarkfin.htm</p> <p><i>"National Archives Guidelines"</i>, UK. http://www.nationalarchives.gov.uk/electronicrecords/advice/guidelines.htm</p>
Datos tipo imagen de documentos de textos, b/n, grises, o color	Periódicos	Formatos escandidos en escala de grises a 400-600 dpi como originales acompañados de pdf y txt para distribución	<p><i>"The National Digital Newspaper Program of the Library of Congress"</i> (NDNP) http://www.loc.gov/ndnp/ndnp_techguide.pdf</p> <p><i>"The british Library Newspaper Pilot"</i>. 2001. http://digitalcooperative.oclc.org/digitize/BritishLibraryNewspaper.html</p>
Imágenes y datos con mapeo de bits escalables al zoom y resolución variada	Mapas, datos cartográficos y geoespaciales, aerofotografía, estereofotografía	TIFF no comprimido o jpeg2000 sin pérdida con paleta y resolución máxima como originales y archivos de menor resolución para distribuir	
Datos con mapeo de bits escalables al zoom y resolución variada	Sistemas de Información Geoespaciales (gis)	Datos alfanuméricos rasterizados o renderizados	<p><i>"gis. A Guide to Good Practice"</i>. http://ads.ahds.ac.uk/project/goodguides/gis/sect11.html</p>

Tipo de Datos	Aplicaciones	Formatos	Referencias
Audio	Archivos de voz y/o Archivos Musicales, tanto restrospectivos como actuales.	Archivos maestros con algún tipo de flujo PCM en formato AIFF, RA, WAV, MPEG1-IV, MPEG1-AAC OMPEG2 no comprimido para preservación. Formatos ligeros como el MPEG1-III (mp3) para distribuir.	<p><i>"Eda Kuhn Loeb MusicLibrary of Harvard College"</i>. http://hcl.harvard.edu/loebmusic/audioreources.html</p> <p><i>"Audio and Video Preservation Reformatting. A Library of Congress Perspective"</i>. http://www.archives.gov/preservation/conferences/papers_2003/fleischhauer.html</p> <p><i>"The Library of Congress Digital Audio Preservation Prototyping Project"</i>. http://www.arl.org/preserv/sound_savings_proceedings/fleischhauer.html</p> <p><i>"The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy. Version 2, September 2001"</i>. (IASA, 2001) http://www.iasa-web.org/iasa0013.htm</p> <p><i>"Digitizing Speech Records for Archival Purposes"</i>. (Historival Voices Project) http://www.historicalvoices.org/papers/audio_digitization.pdf</p> <p><i>"Digitizing Speech Records for Archival Purposes. Digitizing Audio for National Gallery of Spoken Word"</i>. (Historival Voices Project) http://www.historicalvoices.org/research.php</p>
Video	Digitalización de video analógico o cine	MPEG1-IV (MPEG4), RA o QT de alta resolución 1.5-2 Mbps para preservar. Los mismos en baja resolución 28-56 kbps con menos cuadros por seg. y cuadros más pequeños.	<p><i>"The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials"</i>. http://www.nyu.edu/its/humanities/ninchguide/</p> <p><i>"Capture your collections. Small Museum version bibliography"</i>. http://www.chin.gc.ca/English/Digital_Content/Capture_Collections/Publication/bibliography.pdf</p>
Video	Imágenes actuales digitales.	<i>Ídem</i>	<i>Ídem</i>

ALEJANDRÍA,641 a.D.

Jorge Luis Borges

Desde el primer Adán que vio la noche
Y el día y la figura de su mano,
Fabularon los hombres y fijaron
En piedra o en metal o en pergamino
Cuanto ciñe la tierra o plasma el sueño.
Aquí está su labor: la Biblioteca.
Dicen que los volúmenes que abarca
Dejan atrás la cifra de los astros
O de la arena del desierto. El hombre
Que quisiera agotarla perdería
La razón y los ojos temerarios.
Aquí la gran memoria de los siglos
Que fueron, las espadas y los héroes,
Los lacónicos símbolos del álgebra,
El saber que sondea los planetas
Que rigen el destino, las virtudes
De hierbas y marfiles talismánicos,
El verso en que perdura la caricia,
La ciencia que descifra el solitario
Laberinto de Dios, la teología,
La alquimia que en el barro busca el oro
Y las figuraciones del idólatra.
Declaran los infieles que si ardiera,
Ardería la historia. Se equivocan.
Las vigiliass humanas engendraron
Los infinitos libros. Si de todos
No quedara uno solo, volverían
A engendrar cada hoja y cada línea,
Cada trabajo y cada amor de Hércules,
Cada lección de cada manuscrito.
En el siglo primero de la Hégira,
Yo, aquel Omar que sojuzgó a los persas
Y que impone el Islam sobre la tierra,
Ordeno a mis soldados que destruyan
Por el fuego la larga Biblioteca,
Que no perecerá. Loados sean
Dios que no duerme y Muhammad, Su Apóstol.

Bibliotecas y publicaciones digitales. La edición consta de 500 ejemplares. Coordinación editorial, Zindy E. Rodríguez Tamayo. Formación editorial, Ma. Alejandra Romero Ibáñez. Revisión especializada, Francisco Javier González Ortíz. Centro Universitario de Investigaciones Bibliotecológicas/UNAM. Fue impreso en papel cultural ahuesado de 90 gr. en Grupo Edición, S.A. de C.V., ubicado en Xochicalco 619 Col. Vértiz-Narvarte, México, D.F. Se terminó de imprimir en el mes de diciembre de 2006.